Office of Technology Assessment

Congressional Board of the 100th Congress

MORRIS K. UDALL, *Arizona, Chairman*

TED STEVENS, *Alaska, Vice Chairman*

Senate

ORRIN G. HATCH
Utah

CHARLES E. GRASSLEY
Iowa

EDWARD M. KENNEDY
Massachusetts

ERNEST F. HOLLINGS
South Carolina

CLAIBORNE PELL
Rhode Island

House

GEORGE E. BROWN, JR.
California

JOHN D. DINGELL
Michigan

CLARENCE E. MILLER
Ohio

DON SUNDQUIST
Tennessee

AMO HOUGHTON
New York

JOHN H. GIBBONS
(Nonvoting)

Advisory Council

WILLIAM J. PERRY, *Chairman*
H&Q Technology Partners

DAVID S. POTTER, *Vice Chairman*
General Motors Corp. (Ret.)

EARL BEISTLINE
Consultant

CHARLES A. BOWSHER
General Accounting Office

S. DAVID FREEMAN
Lower Colorado River Authority

MICHEL T. HALBOUTY
Michel T. Halbouty Energy Co.

NEIL E. HARL
Iowa State University

JAMES C. HUNT
University of Tennessee

JOSHUA LEDERBERG
Rockefeller University

CHASE N. PETERSON
University of Utah

SALLY RIDE
Stanford University

JOSEPH E. ROSS
Congressional Research Service

Director

JOHN H. GIBBONS

The Technology Assessment Board approves the release of this report. The views expressed in this report are not necessarily those of the Board, OTA Advisory Council, or individual members thereof.

COVER
Photo credit: Sibley Memorial Hospital

The Quality of Medical Care

Information for Consumers

CONGRESS OF THE UNITED STATES OFFICE OF TECHNOLOGY ASSESSMENT

J952-55

Recommended Citation:

U.S. Congress, Office of Technology Assessment, *The Quality of Medical Care: Information for Consumers*, OTA-H-386 (Washington, DC: U.S. Government Printing Office, June 1988).

Library of Congress Catalog Card Number 88-600537

For sale by the Superintendent of Documents
U.S. Government Printing Office, Washington, DC 20402-9325
(order form can be found in the back of this report)

Foreword

For quite some time, people within the medical profession have been concerned about assessing the quality of medical care so that providers could improve it. Florence Nightingale in the field hospitals of the Crimean War and Ernest A. Codman in Boston's surgical wards during the early 20th century were part of this tradition. Although experts from other fields, such as statistics, contributed techniques to evaluate the quality of medical care, until lately assessments of quality remained largely within the purview of the medical profession.

In recent years, a number of forces have combined to promote consumers' role in evaluating medical providers. Efforts to advance consumers' interests are occurring throughout society, and changes within medical care are part of that societal trend. More specific to medical care are changes in policies designed to inject greater price competition into medical care. According to competitive theory, consumers who are sensitive to both price and quality will bring these considerations to bear as they select health insurance and medical providers. Changes in how physicians and hospitals are paid have made individual consumers, health insurers, employers, and medical providers more sensitive to the cost implications of their decisions. At the same time, these policy changes have elevated the importance of having consumers be informed about the quality of medical providers. Purchasers of medical care (individual consumers, employers, health insurers) need to know about any differences in quality so that they can weigh quality along with cost in making decisions. Furthermore, payment changes have raised the concern that physicians and hospitals facing restricted budgets and low payment rates will skimp on the services that they provide to the detriment of their patients' health.

Congressional interest in public information on the quality of medical care predated the new policies, but these payment changes, especially within the Medicare program, have heightened that interest. It was in that context that the House Committee on Energy and Commerce and its Subcommittee on Health and the Environment requested the Office of Technology Assessment (OTA) to assess whether valid information could be developed and disseminated to the public to assist their choices of physicians and hospitals. The Senate Committee on Finance asked that OTA address several issues related to the availability and confidentiality of data that could be used to assess the quality of medical care. The Senate Select Committee on Aging; the Subcommittee on Consumer of the Senate Committee on Commerce, Science, and Transportation; and the House Committee on Science, Space, and Technology also endorsed the study.

In preparing this report, OTA staff drew on the expertise of members of the advisory panel, chaired by Dr. Frederick Mosteller, and experts in consumer advocacy, medical practice, health insurance, rural health, and quality assessment. Drafts of the report were reviewed by the advisory panel and by numerous individuals and organizations with expertise and interest in the area. We are grateful for their assistance. Key OTA staff for this analysis were Jane E. Sisk, Denise Dougherty, Polly M. Ehrenhaft, Mark McClellan, Beth A. Mitchner, Gloria Ruby, and Kerry Britten Kemp.

John H Gibbons

JOHN H. GIBBONS
Director

Advisory Panel—The Quality of Medical Care: Information for Consumers

Frederick Mosteller, *Panel Chair*
Professor Emeritus, Department of Health Policy and Management
Harvard School of Public Health, Boston, MA

Linda H. Aiken
Trustee Professor of Nursing and Sociology
University of Pennsylvania
Philadelphia, PA

Paul Batalden
Vice President for Medical Care
Hospital Corporation of America
Nashville, TN

Donald Berwick
Vice President, Quality of Care Measurements
Harvard Community Health Plan
Boston, MA

Robert Brook
Deputy Director for Health Services
The Rand Corporation
Santa Monica, CA

Avedis Donabedian
Nathan Sinai Distinguished Professor of Public Health, School of Public Health
The University of Michigan
Ann Arbor, MI

Barbara Herzog
Director, AARP Health Care Campaign
American Association of Retired Persons
Washington, DC

David Kanouse
Senior Behavioral Scientist, The Rand Corp.
Santa Monica, CA

Gerald C. Kempthorne
Medical Director, HMO of Wisconsin
Spring Green, WI

Kathleen N. Lohr
Senior Professional Associate
Institute of Medicine
Washington, DC

Floyd Loop
Chair, Department of Thoracic and Cardiovascular Surgery
Cleveland Clinic Foundation
Cleveland, OH

Lorna McBarnette
Executive Deputy Commissioner of Health
New York State Department of Health
Albany, NY

William H. Moncrief, Jr.
President, California Medical Review, Inc.
San Francisco, CA

Peter O'Donnell
Senior Vice President
Alta Health Strategies, Inc.
Princeton, NJ

R. Heather Palmer
Lecturer in Health Services
Harvard School of Public Health and Institute for Health Research
Boston, MA

James S. Roberts
Vice President for Accreditation
Joint Commission on Accreditation of Healthcare Organizations
Chicago, IL

Ruth Ruttenberg
President, Ruth Ruttenberg Associates, Inc.
Bethesda, MD

Cathy Schoen
Research Director
Service Employees International Union
Washington, DC

Laurence R. Tancredi
Director, Health Law Program
University of Texas
Health Sciences Center at Houston
Houston, TX

Sidney Wolfe
Director, Public Citizen Health Research Group
Washington, DC

James M. Young
Lecturer, Harvard School of Public Health
Boston, MA

NOTE: OTA gratefully acknowledges the members of this advisory panel for their valuable assistance and thoughtful advice. The panel does not, however, necessarily approve, disapprove, or endorse this report. OTA assumes full responsibility for the report and the accuracy of its contents.

OTA Project Staff—The Quality of Medical Care: Information for Consumers

Roger C. Herdman, *Assistant Director, OTA*
Health and Life Sciences Division

Clyde J. Behney, *Health Program Manager*

Project Staff

Jane E. Sisk, *Project Director*

Denise M. Dougherty, *Analyst*

Polly M. Ehrenhaft, *Senior Analyst*

Mark McClellan, *Research Assistant*[*]

Beth A. Mitchner, *Research Analyst*

Gloria Ruby, *Senior Analyst*

Kerry Britten Kemp, *Division Editor/Analyst*

Katherine Eddy Cox, *Research Assistant*[**]

Administrative Staff

Virginia Cwalina, *Administrative Assistant*

Carol Ann Guntow, *P.C. Specialist*

Karen T. Davis, *Secretary/Word Processor Specialist*

Carolyn Martin, *Clerical Assistant*

Major Contractors

Nancy E. Cahill, *Duke University, North Carolina*

Peter G. Goldschmidt, *World Development Group, Inc.*

Karen Glanz, Joel Rudd, *University of Minnesota, University of Arizona*

Marlene Larks, *National Association of Health Data Organizations*

Harold S. Luft, Deborah W. Garnick, David Mark, Stephen J. McPhee, Janice Tetreault,
University of California, San Francisco

Don Harper Mills, Orley Lindgren, *Institute for Medical Risk Studies*

James B. Simpson, *Western Consortium for the Health Professions, Inc.*

John E. Ware, Jr., Allyson Ross Davies, Haya H. Rubin, *The Rand Corp.*

---

[*] Summer 1987; under contract February through March 1988.

[**] From March 1988.

Contents

Glossary of Abbreviations and Terms vii

Chapter *Page*

1. Summary and Policy Implications 3
2. Disseminating Information to Consumers: Present Context and Future Strategy 33
3. Evaluating Quality From the Perspective of Individual Consumers 51
4. Hospital Mortality Rates 71
5. Adverse Events 101
6. Disciplinary Actions, Sanctions, and Malpractice Compensation 121
7. Evaluation of Physicians' Performance: Care for Hypertension 145
8. Volume of Services in Hospitals or Performed by Physicians 165
9. Scope of Hospital Services: External Standards and Guidelines 187
10. Physician Specialization 209
11. Patients' Assessments of Their Care 231

Appendixes *Page*

A. Method of the Study 251

B. Acknowledgments 253

C. Method Used by OTA To Evaluate Indicators of Quality 257

D. Quality Assessment Activities by Selected Organizations 274

E. Selected Studies Related to the Quality of Medical Care 286

References 291

Glossary of Abbreviations and Terms

Glossary of Abbreviations

ABMS	—American Board of Medical Specialties
ACEP	—American College of Emergency Physicians
ACOG	—American College of Obstetricians and Gynecologists
AMA	—American Medical Association
APACHE	—Acute Physiology and Chronic Health Evaluation
CABG	—coronary artery bypass graft surgery
CDC	—Centers for Disease Control (HHS)
CFR	—Code of Federal Regulations
CPHA	—Commission on Professional and Hospital Activities
CMP	—competitive medical plan
COBRA	—Consolidated Omnibus Budget Reconciliation Act of 1985 (Public Law 99-272)
DHHS	—U.S. Department of Health and Human Services
DRG	—diagnosis-related group
ENA	—Emergency Nurses Association
HCFA	—Health Care Financing Administration (HHS)
HHS	—U.S. Department of Health and Human Services
HMO	—health maintenance organization
ICD-9-CM	—International Classification of Diseases, Ninth Revision, Clinical Modification
IPA	—individual practice association
JCAHO	—Joint Commission on the Accreditation of Healthcare Organizations
MEDISGRPS	—Medical Illness Severity Grouping System
OIG	—Office of the Inspector General (HHS)
OBRA-86	—Omnibus Budget Reconciliation Act of 1986 (Public Law 99-509)
OBRA-87	—Omnibus Budget Reconciliation Act of 1987 (Public Law 100-203)
OTA	—Office of Technology Assessment (U.S. Congress)
PPI	—Physician Performance Index
PPS	—prospective payment system
PRO	—utilization and quality control peer review organization
SENIC	—Study on the Efficacy of Nosocomial Infection Control (CDC study)
VA	—Veterans Administration

Glossary of Terms

Access: Potential and actual entry of a population into the health care delivery system.

Accreditation by JCAHO: A statement by the Joint Commission on the Accreditation of Healthcare Organizations that an eligible health care organization, such as a hospital, complies wholly or substantially with JCAHO standards. Hospitals or other health care organizations that are surveyed but do not meet JCAHO standards are referred to as *nonaccredited*. Hospitals that either do not request a survey or are not eligible to be surveyed are referred to as *unaccredited*. Compare *certification by HCFA*.

Acute myocardial infarction: Necrosis (death) of tissue in the myocardium (heart muscle) that results from insufficient blood supply to the heart.

Adverse events: Untoward events involving patients. Adverse events are typically unanticipated poor patient outcomes, such as death or readmission to the hospital. Other incidents such as improper administration of medications or patient falls are also considered adverse events even if there is no effect on the patient. See *incident reporting* and *occurrence screen*.

Ambulatory care: Medical services provided to patients who have not been admitted to a hospital or nursing home.

Aneurysm: A permanent, abnormal, blood-filled dilation of a blood vessel or the heart resulting from disease of the vessel or heart wall.

APACHE: A system that uses physiological values, age, and certain aspects of chronic health status to measure a patient's risk of dying. The system has been applied chiefly to patients in hospital intensive care units.

Bacteremia: The presence of bacteria in the blood.

Biliary tract surgery: Surgery involving the bile-conveying structures (duodenum, gall bladder, liver).

Board certification: A method of formally identifying a physician who has completed a specified amount of training and a certain set of requirements, and passed an examination required by a medical specialty board.

Cardiac catheterization: The passage of a catheter through a vein into the heart for diagnostic purposes.

Case finding: The identification of instances of a particular disease or condition through screening of asymptomatic people or surveillance of defined populations.

Case mix: The relative frequency of different medical conditions or diagnoses among patients.

Certification by HCFA: A statement by the Health Care Financing Administration (HCFA) that a hospital meets HCFA's conditions of participation. Certification by HCFA is required for Medicare and Medicaid reimbursement. Compare *accreditation by JCAHO*.

Certification by a medical specialty board: See *board certification*.

Cholecystectomy: Surgical removal of the gall bladder.

Claims data: Data derived from medical providers' claims to third-party payers.

Clinical data: Data on patients derived from clinical examination and tests.

Comorbidities: Diseases or conditions present at the same time as the principal condition of a patient.

Complications: Adverse patient conditions that arise during the process of medical care.

Contingency: A decision by the Joint Commission on the Accreditation of Healthcare Organizations (JCAHO) that a hospital is in substantial noncompliance with the requirements for a certain JCAHO standard. The hospital must then conform to that standard within a time period that is shorter than the 3-year accreditation cycle, or risk nonaccreditation.

Coronary artery bypass graft (CABG) surgery: A surgical procedure in which a vein or an artery is used to bypass a constricted portion of one or more coronary arteries. This procedure has become the primary surgical approach to the treatment of coronary artery disease.

Diagnosis-related groups (DRGs): Groupings of diagnostic categories drawn from the International Classification of Diseases and modified by the presence of a surgical procedure, patient age, presence or absence of significant comorbidities or complications, and other relevant criteria. DRGs are the case-mix measure mandated for Medicare's prospective hospital payment system by the Social Security Amendments of 1983 (Public Law 98-21).

Discharge abstract: A summary of data abstracted from a hospitalized patient's medical record that usually includes specific clinical data such as diagnostic and procedure codes as well as other information about the patient, the physician, and insurance and financial status.

Disciplinary actions by State medical boards: See *State medical boards' disciplinary actions*.

Efficacy: The probability of benefit to individuals in a defined population from a medical technology applied for a given medical problem under ideal conditions of use.

Explicit review: Review of the process of medical care using explicit criteria specified in advance. Compare *implicit review*.

External validity: See *validity*.

Face validity: See *validity*.

False negative: A negative result in a case that actually has the condition or characteristic for which a test was conducted.

False positive: A positive result in a case that does not have the condition or characteristic for which a test was conducted.

Feasibility: In the context of evaluations of indicators of medical quality, whether it is practical to use a certain indicator to convey information to the public about quality.

Femur fracture: Fracture of the thigh bone.

Generic screen: See *HCFA generic quality screens*.

Gross and flagrant violation: A violation that presents an imminent danger to the health, safety, or wellbeing of a Medicare beneficiary or that unnecessarily places the beneficiary at risk of substantial and permanent harm. Utilization and quality control peer review orgranizations (PROs) identify potential violations and recommend sanctions, but the Office of the Inspector General of the U.S. Department of Health and Human Services makes the final decision as to whether to impose sanctions. Compare *substantial violation*.

HCFA generic quality screens: The list of occurrences applied by utilization and quality control peer review organizations (PROs) to select cases that may have quality problems and that merit scrutiny. Because these screens generate a large portion of false positives, their application is only the first step in a multistage review process.

Health maintenance organization (HMO): A health care organization that, in return for prospective per capita (capitation) payments, acts as both insurer and provider of comprehensive but specified medical services. A defined set of physicians provide services to a voluntarily enrolled population. Prepaid group practices and individual practice associations are types of HMOs.

Hernia: Any abnormal protrusion of one anatomical structure through another. The most common variety is herniation of part of the intestine through a weakness in the abdominal wall.

High-mortality outliers: Providers with mortality rates that are higher than expected after adjustment for patient or other characteristics. Compare *low-mortality outliers*.

Hospital accreditation: See *accreditation by JCAHO*.

Hospital discharge abstract: See *discharge abstract*.

Hospital mortality rate: Number of deaths as a proportion of the total number of hospital patients or admissions. See *mortality rate*.

Hospital volume: The number of a particular procedure performed or condition treated in a hospital. See *volume*.

Hypertension: Persistently high blood pressure. The chief importance of hypertension lies in the increased risk it confers of illness and death from cardiovascular, cerebrovascular, and renal disease.

Hysterectomy: Surgical removal of the uterus.

Iatrogenic illness: Any adverse condition in a patient that is caused by medical treatment.

Impaired physician: A physician who does not have the ability to practice medicine with reasonable skill and safety to patients because of physical or mental illness, including alcoholism or drug dependence.

Implicit review: Review of the process of medical care using subjective criteria. Compare *explicit review*.

Incidence: The frequency of new occurrences of a condition within a defined time interval. Incidence rate is the number of new cases of specified disease divided by the number of people in a population over a specified period of time, usually 1 year. Compare *prevalence*.

Incident reporting: A system for collecting and reporting information about adverse events that affect patients in hospitals. Hospital personnel (most frequently nurses) complete forms when they observe an adverse event; the definition of an "incident" is discretionary by the frontline health professionals who deal with patients. Examples of incidents include patient falls, medication errors, equipment failures, and procedure or treatment errors.

Inpatient care: Medical services provided to patients who have been admitted to hospitals.

Internal validity: See *validity*.

International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) Coding: A two-part system of coding patient medical information used in abstracting systems and for classifying patients into DRGs for Medicare. The first part is a comprehensive list of diseases with corresponding codes compatible with the World Health Organization's list of disease codes. The second part contains procedure codes, independent of the disease codes.

Interpersonal aspects of medical care: The personal interaction between patient and provider.

Interrater reliability: Consistency of judgments among raters or sets of raters.

Intrarater reliability: Consistency of judgments by a single rater.

Liability: Accountability and responsibility that are enforceable by legal sanctions.

Licensure: The process by which a State grants permission to a physician to practice medicine upon finding that she or he has met acceptable qualification standards. Licensure also involves ongoing State regulation of physicians, including the State's authority to revoke or otherwise restrict a physician's license to practice.

Low-mortality outliers: Providers with mortality rates that are higher than expected after adjustment for patient or other characteristics. Compare *high-mortality outliers*.

Medicaid: A federally aided, State-administered program that provides medical assistance to certain low-income people.

Medical injury: An adverse outcome that could be either unavoidable or avoidable, i.e., negligently induced.

Medical malpractice: A judicial determination that there has been a negligent (or, rarely, willful) failure to adhere to the current standards of medical care, resulting in injury to the patient. Since the judgment of malpractice is social-legal and is made on a case-by-case rather than systematic basis, standards and processes for determining malpractice vary by area.

Medical practice act: A State law that provides statutory authority for the State to license and discipline physicians and other health care professionals.

Medical record: The account compiled by physicians or other medical professionals of patients' medical history, present illness, findings on examination, details of treatment, and notes on progress. The medical record is the legal record of care.

Medical record audit: See *medical record review*.

Medical record review: Review of a patient's medical record to determine how the medical provider performed.

Medical technology: The drugs, devices, and medical and surgical procedures used in medical care, and the organizational and support systems within which such care is provided.

Medicare: A nationwide, federally administered health insurance program first authorized in 1965 that now covers hospitalization, physician care, and some related services for eligible persons over age 65, persons receiving Social Security Disability Insurance payments for 2 years, and persons with end-stage renal disease.

Medicare conditions of participation: Requirements that institutional providers (including hospitals, skilled nursing homes, home health agencies, etc.) must meet in order to be allowed to receive payments for Medicare patients. An example is the requirement that hospitals conduct utilization review.

Medical Illness Severity Grouping System (MEDISGRPS): A computerized data system developed by MediQual Systems, Inc., that categorizes patients' risk of dying or of increased morbidity based on key physiological findings.

MEDLINE data base: The original, largest, and most utilized data base in the National Library of Medicine's computerized retrieval and technical processing system. MEDLINE contains references to biomedical and other literature relevant to health and health services.

Meta-analysis: The quantitative analysis of a large collection of results from individual studies for the purpose of integrating the findings.

Morbidity rate: The rate of illness in a population. The number of people ill during a time period divided by the number of people in the total population.

Mortality rate: The death rate, often made explicit for a particular characteristic, e.g., age, sex, or specific cause of death. A mortality rate contains three essential elements: 1) the number of people in a population group exposed to the risk of death (the denominator); 2) a time factor; and 3) the number of deaths occurring in the exposed population during a certain time period (the numerator).

Negotiated settlement: The resolution of a malpractice claim prior to a judicial determination.

Neonatal: Pertaining to the first 4 weeks after birth.

Nonaccreditation: See *accreditation by JCAHO*.

Nosocomial infection: An infection that a patient acquires in a hospital or other institution. The most common nosocomial infections are urinary tract infections, followed by surgical wound infections, pneumonia, and infections of the bloodstream.

Occurrences: Adverse events. See *adverse events*.

Occurrence screen: A list of criteria used to screen patients' medical records for occurrences. Examples of occurrences include deaths, unusually long lengths of stay, hospital-acquired infections, and unscheduled procedures, readmissions, or transfers.

Outcome measures of quality: Measures of changes in patient outcomes, that is, patient health status and satisfaction. Attributing changes in outcomes to medical care requires distinguishing the effects of care from the effects of the many other factors that influence patients' health and satisfaction.

Outliers: See *high-mortality outliers* and *low-mortality outliers*.

Outpatient care: Care that is provided in a hospital and that does not include an overnight stay.

p value: The probability of concluding that a statistical association exists between, for instance, a risk factor and a health endpoint, when, in fact, there is no real association. In other words, the likelihood that an observed association in a study is due to chance. Also called "Type I error" or "alpha," and commonly called the level of significance.

Patients' assessments: Patients' ratings and reports.

Patients' ratings: Personal evaluations of aspects of medical care providers and services. Ratings are inherently subjective because they reflect personal experiences, expectations, and preferences, as well as the standards patients apply when evaluating care. Compare *patients' reports*.

Patients' reports: Information from patients about things that did or did not happen during their medical care. Patients' reports are inherently more objective than patients' ratings and can be more readily confirmed by an outside observer. Compare *patients' ratings*.

Peer review organizations: See *utilization and quality control peer review organizations*.

Perinatal: Pertaining to or occurring in the period shortly before and after birth; variously defined as beginning with the completion of the 20th to 28th week of gestation and ending 7 to 28 days after birth.

Physician credentialing: A process that includes education, licensure, specialty certification, and conferring hospital privileges and that is intended to ensure physician competence and protect public safety.

Physician Performance Index (PPI): A process measure of physician performance that evaluates physicians' compliance with certain explicit criteria. The criteria were weighted by a panel of physicians and aggregated to generate a single PPI score for each diagnosis or examination. A physician performance score represents a physician's average PPI score over all of his or her treated cases.

Physician volume: The number of a procedure performed or condition treated by individual physicians. See *volume*.

Predictive validity: See *validity*.

Prevalence: The number of existing cases of a disease or condition in a given population at a specific time. Compare *incidence*.

Principal diagnosis: The diagnosis which, after study, is judged to be the principal reason for hospitalization or other medical care.

Process measures of quality: Measures of the activities of physicians and other health professionals in caring for patients. To evaluate providers' performance, it is valid to use only process measures that have been shown to improve or harm patients' health and satisfaction, a link that has been established for relatively few processes.

Prospective study: A study in which data are gathered after a hypothesis has been generated and the study approved. Compare *retrospective study*.

Prostatectomy: Surgical removal of the prostate gland.

Quality of medical care: Evaluation of the performance of medical providers according to the degree to which the process of care increases the probability of outcomes desired by patients and reduces the probability of undesired outcomes, given the state of medical knowledge. Which elements of patient outcomes predominate depends on the patient condition.

Quality assessment: Measurement and evaluation of the quality of medical care for individuals, groups, or populations.

Quality assurance: Activities to safeguard or improve the quality of medical care by assessing quality and taking action to correct any problems found.

Randomized trial: A study in which subjects are assigned randomly to either the experimental or the control condition.

Readmission: Admission to a hospital within a specified period of time after a prior admission or because of complications of a prior admission.

Regression analysis: A statistical procedure for determining the best approximation of the relationship between variables. Multiple regression analysis is a method for measuring the effects of several factors concurrently.

Reliability: Consistency in results of a measure, including the tendency of the measurement to produce the same results twice when it measures some entity or attribute believed not to have changed in the interval of measurements. Reliability is a prerequisite to validity. See *interrater reliability* and *intrarater reliability*.

Retrospective study: A study in which data that are already available are analyzed to test a hypothesis. Compare *prospective study*.

Risk management: Programs that institutions, especially hospitals, undertake to prevent medical mishaps and to minimize the adverse effects of injury and loss to patients, employees, visitors, and the institution itself. Quality assurance is often considered a subset of the larger issue of risk management.

Scope of hospital services: A structural measure of the quality of care that reflects whether a hospital has the resources—facilities, staff, and equipment—to provide care for the medical conditions it professes to treat or to care for the medical condition affecting a potential patient.

Secondary diagnosis: Any medical condition of a patient other than the principal diagnosis. See *comorbidities*.

Selective referral: The referral or attraction of patients to physicians and hospitals with better outcomes.

Sensitivity of a test: For a particular test, the percentage of individuals who actually have the condition being tested for who are correctly identified as positive by the test. Operationally, sensitivity is the number of true positive test results divided by the number that actually have the condition (true positives divided by the sum of true positives plus false negatives). Compare *specificity of a test*.

Specificity of a test: For a particular test, the percentage of individuals who do not have the condition being tested for who are correctly identified as negative by the test. Operationally, specificity is the number of negative test results divided by the number of individuals who actually do not have the condition (true negatives divided by the sum of true negatives plus false positives). Compare *sensitivity of a test*.

State medical boards: State licensing bodies and State disciplinary bodies. States exercise their authority to license physicians through State licensing boards. The disciplinary functions may be incorporated in the same body as the licensure function or in a separate body.

State medical boards' disciplinary actions: The penalties imposed by State medical boards on physicians who have transgressed provisions in State medical practice acts. The penalties range from revoking licenses to practice medicine through lesser penalties such as suspension of licenses for a period of time; probation; stipulations; limitations and conditions relating to practice; reprimands; letters of censure and letters of concern.

Statistical conclusion validity: The extent to which research is sufficiently precise or powerful to enable observers to detect effects. Conclusion errors are of two types: Type I is to conclude there are effects (or relationships) when there are not; Type II is to conclude there are no effects (or relationships) when in fact they exist.

Statistical power: The probability of detecting a difference between the groups being compared when one does exist. Failure to detect an effect is called "Type II error" or "beta," analogous to "false negative."

Statistically significant: The likelihood that an observed association is not due to chance. See *p value*.

Structural measures of quality: Measures of the resources and organizational arrangements that are in place to deliver medical care, such as the number, type, and distribution of medical personnel, equipment, and facilities. Underlying the use of such measures to assess quality is the assumption that such characteristics increase or decrease the likelihood that providers will perform well and their absence, that providers will perform poorly. This assumption in turn raises the question whether specific structural characteristics are in fact associated with better process or outcome.

Substantial violation: A pattern of care over a substantial number of cases that is inappropriate, unnecessary, does not meet the recognized standards of care, or is not supported by the documentation of care required by the PRO. PROs identify potential violations; the Office of the Inspector General of the U.S. Department of Health and Human Services makes the final decision as to whether the violation occurred. Compare *gross and flagrant violation*.

SuperPRO: An independent organization, working under contract to HCFA, that re-reviews a sample of the patient records evaluated by each of the 54 PROs. The purpose of the SuperPRO reviews is to validate the determinations made by PROs, including the application of the HCFA generic quality screens. To date, the SuperPRO contract has been held by SysteMetrics, Inc., in Santa Barbara, California.

Targeted mortality method: An approach to quality assessment used, for example, by the New York State Department of Health in which deaths in certain types of cases are targeted for review. Examples include deaths in primary procedures or DRGs with an average death rate of less than 5 percent, deaths occurring within 1 day of any procedure, and deaths in which burns are reported as a secondary diagnosis.

Technical aspects of medical care: The application of medical science and technology to a medical problem.

Third-party payment: Payment by a private insurer or government program to a medical provider for care given to a patient.

Tort liability: Liability imposed by a court for breach of a duty implied by law, contrasted with contractual liability, which is breach of duty arising from an agreement. The tort liability system determines fault and awards compensation for civil wrongs, including medical malpractice.

UB-82: The uniform billing form required by the Health Care Financing Administration for submitting and processing Medicare claims. It merges billing information with diagnostic codes, including almost all the elements from the uniform hospital discharge data set.

Unaccredited: See *accreditation by JCAHO*.

Utilization and quality control peer review organizations (PROs): Organizations established by the Tax Equity and Fiscal Responsibility Act of 1982 (Public Law 97-248) with which the U.S. Department of Health and Human Services contracts to review the appropriateness of settings of care and the quality of care provided to Medicare beneficiaries.

Validity: A measure of the extent to which an observed situation reflects the true situation or an indicator of medical quality measures what it purports to measure. There are several types of validity:

Construct validity: The extent to which an indicator measures what it is supposed to measure. If construct validity has been established for a measure, it may be used as a criterion or gold standard against which other measures (tests, indicators) are evaluated.

Content validity: How representative a sample of items is of the universe that it was intended to represent.

Convergent validity: A demonstration of the validity of a measure by correlations among two or more purported measures of a concept. Convergent validity does not, however, presuppose that one measure is a standard against which other measures should be evaluated.

Discriminant validity: A demonstration of the validity of a measure by the lack of correlation among two or more supposedly unrelated measures of a concept.

External validity: The extent to which the results of a study may be generalized beyond the subjects of a study to other settings, providers, procedures, diagnoses, etc.

Face validity: Intelligibility, i.e., the extent to which an indicator and hypothesized relationships would make sense to the average consumer and provider.

Internal validity: The extent to which the design of a study contributes to the confidence that can be placed in the study's results. Internal validity is relevant to both measurement studies and studies of causal relationships; it is the extent to which the detected relationships are most likely due to factors accounted for in the study, rather than other factors.

Predictive validity: The ability of an indicator to predict future events.

Validity variable: A measure of the quality of care derived independently of the indicator being evaluated. Experimental manipulation, observation and simulation, provider report, or chart review are all sources of information for validity variables.

Volume: The number of cases with a specific procedure, diagnosis, or condition treated in a hospital or by a physician.

Chapter 1

Summary and Policy Options

CONTENTS

Page

Introduction 3
Scope of the Study 4
Summary 7
  Quality From the Perspective of Individual Consumers 7
  Findings Regarding Specific Indicators of Quality 10
Policy Options 17
  To Improve Quality Assessment Techniques 19
  To Ensure the Quality of Quality Assessments 20
  To Improve the Availability of Required Data 23
  To Disclose Information to the Public 25
  To Disseminate Information to the Public 28
Conclusions 30

Tables

Table *Page*

1-1. Indicators of the Quality of Care Evaluated by OTA, by Type of Medical Provider 6
1-2. Indicators of the Quality of Care Evaluated by OTA, by Assessment Approach 8
1-3. Indicators of the Quality of Care Evaluated by OTA, by Aspects of Medical Care 11
1-4. Summary of Key Findings on Quality-of-Care Indicators Evaluated by OTA 14
1-5. Summary of Policy Options for Congress To Address Problems With Quality-of-Care Indicators 18

Chapter 1

Summary and Policy Options

INTRODUCTION

In the wake of recent policies to contain medical expenditures has come a ground swell of support for public information on the quality of individual medical providers. The call for better information comes from many quarters—policymakers, consumer advocates, large-scale purchasers of medical care, and medical professionals, all groups with a longstanding interest in the caliber of medical care. For quite some time, payment policies that reward the use of extra services and expensive procedures have posed a threat to the quality of care by creating incentives to provide care that may be inappropriate. But recent changes in payment policies have raised concerns from another direction—that hospitals and physicians facing restricted budgets and low payment rates will skimp on services to the detriment of patients' health and that third-party payers will seek low-cost providers with insufficient regard for their quality of care.

In the present environment, at least three rationales lie behind the call for more public information on the quality of medical providers. The most immediate is that people seeking medical care deserve information so that they can avoid poor providers and seek good providers. This rationale assumes that some medical providers may harm patients or may furnish care much inferior to that of other providers. The second rationale for more public information is that over a longer period of time, information on specific providers could form part of a larger effort to educate the public about the quality of medical care. Indeed, informed consumers play a pivotal role in strategies to inject greater price competition into the medical marketplace. According to competitive theory, the decisions of consumers weighing price and quality levels and selecting health insurance and medical providers guide the cost and quality of care that result. As payment changes have made individual consumers, their agents, and medical providers more sensitive to price, it has grown even more important that purchasers of medical care (individual consumers, employers, and third-party payers) know about any differences in the quality of care. Only with information about quality will people making decisions be able to weigh quality along with cost. A general educational effort could impart the knowledge and skills to enable people to appreciate differences in the quality of care offered by medical providers.

Photo credit: March of Dimes Birth Defects Foundation

The call for better information on the quality of hospitals and physicians has come from consumer advocates, large-scale purchasers of medical care, medical professionals, and policymakers.

A third rationale for better public information on the quality of care is to stimulate the medical community, as a collective and as individuals, to improve their quality. From the choices of informed purchasers, medical providers can gain insight into what matters to people who seek medical care. Some policymakers and medical professionals envisage that the increased knowledge from such feedback and the competition for patients will drive medical providers, both hospitals and physicians, to better their own practices.

The current focus on the quality of care needs to be put into the broader context of U.S. medicine. The U.S. medical delivery system has made enormous advances in the health of the Nation, some to lengthen life and others to improve its quality. Perhaps the very successes of U.S. medicine have spawned the calls for more quality assessment and public information, for along with these achievements, public expectations of medicine and the public's stake in good-quality care have risen. People now have much more to gain from medicine, and much more to lose from poor-quality care. At the same time, several studies have found much room for improvement among different types of providers and disturbing variations in the use of medical procedures and hospital care (79,131,215,696). Furthermore, improvements in health have not been uniform or universal, and some people, notably the underinsured and uninsured, receive less care than others.

Congress has long had an interest in public information on medical care, especially as it relates to the Medicare program. In recent years, changes in payment have heightened that interest, as public and private payers have adopted policies intended to increase price competition in medical care. In October 1983, for example, Medicare changed its system of payment for inpatient operating expenses to a system of payments set in advance and varying according to the patient's diagnosis-related group (DRG) (630). Medicare's present payment system gives hospitals an incentive to be frugal about aspects of care that add to their operating costs without adding to their revenue. Sizable reductions in Medicare beneficiaries' lengths of hospital stay and days in intensive care units suggest that medical providers are in fact trimming resource use (620).

Reducing hospitalized patients' lengths of stay and intensity of resource use may improve the patients' health and the quality of care to the extent that nosocomial (hospital-acquired) infections and iatrogenic (medically caused) problems are avoided, that more extensive technology use carries some risk and adds little or nothing to patients' health, and that a shorter stay or lower level of care is equally or more appropriate. On the other hand, the quality of care may be impaired if tests and procedures that would benefit patients' health are not used, if earlier hospital discharge and care at a lower level harm patients' health, or if delay in more intensive treatment jeopardizes patients' conditions. Certain populations are especially vulnerable to the effects of public and private cost containment: poor people because they are more dependent on public programs for their care; severely ill people because providers may wish to avoid their admission, transfer them, or discharge them early; and physically or mentally impaired people because they have less ability to cope with the system.

In this context, the House Committee on Energy and Commerce and its Subcommittee on Health and the Environment requested the Office of Technology Assessment (OTA) to assess whether information could be developed and distributed to the public to assist their choice of medical providers. The committee asked whether there were valid indicators of the quality of care that consumers could use to select physicians and acute-care hospitals. In addition, the Senate Committee on Finance; the Senate Select Committee on Aging; the Subcommittee on Consumer of the Senate Committee on Commerce, Science, and Transportation; and the House Committee on Science, Space, and Technology endorsed the study. The Senate Committee on Finance asked that OTA specifically address several issues related to data, including their availability, confidentiality, and access. This report responds to the requests of those committees.

SCOPE OF THE STUDY

This OTA report evaluates the reliability, validity, and feasibility of specific indicators of the quality of medical care that purchasers of care—individuals, employers, and third-party payers—might use. Reflecting the committees' interests and OTA's time constraints, the report deals with indicators of quality only for physicians and acute-care hospitals. Although the quality of health insurance plans lies beyond the scope of this study, the conclusions of the study apply to hospitals and physicians affiliated with such plans, including health maintenance organizations (HMOs) and preferred provider organizations (PPOs). Given the report's focus on physicians and acute-care hospitals, the report also excludes indicators of quality for medical professionals other than physicians and for providers of long-term care, such as nursing homes and home care agencies. Nevertheless, these topics merit attention as policymakers consider consumer choice and public disclosure of information. For most Americans, which physicians and hospitals are financially accessible hinges on health insurance coverage. Within hospitals and other organizations, the quality of care depends not only on physicians, but also on nurses and other health professionals and on coordination among different health professionals (570). The importance of the quality of long-term care has mushroomed as constraints on hospitals have restricted admissions and spurred earlier discharges.

As a result of limiting the analysis to hospitals and physicians, the report considers how to evaluate the care received by people who seek care and receive services, but does not consider how to evaluate the quality of the entire U.S. health care system. Most issues relating to the accessibility of care to individuals are thus excluded from this report. Numerous factors—psychological, physical, social, and economic—determine whether a person seeks care for a medical condition. Among them is the cost that the person expects to pay, which in turn depends on insurance coverage (or the lack of it) and the provider's charges. Most hospitals and physicians practice independently and do not assume responsibility for ensuring that certain services are available to a clearly defined population. It would not be reasonable to hold these providers responsible for the ease of access by all the people in an area. Once an individual has established a relationship with a provider, however, it seems reasonable to hold the provider responsible for making medical services accessible to those patients.

Also excluded from this report are considerations of cost and efficiency. Medical costs indicate what people forgo in other goods and services to obtain the health outcomes that they desire. In making decisions about medical care, purchasers weigh the likely costs and benefits, as they do for other goods and services. In fact, behind many of the recent changes in payment policies has lain the intention of heightening the cost consciousness of consumers and providers about using medical services. Although decisionmaking requires consideration of both cost and quality, separating issues of cost and quality reflects that health effects are distinct and that costs are incurred to obtain the health effects desired.

Technology assessment should undergird assessment of the quality of a provider's practice (103). Using standards to evaluate the quality of care delivered to a patient requires that a quality assessor have criteria by which to judge how a particular condition is managed. The development of such criteria, in turn, should be based on knowledge about the efficacy and safety of new and existing medical technologies. Thus, quality assessment requires information from prior technology assessments about the benefits and risks of technologies under routine and ideal conditions of use. For a given technology, an initial technology assessment is unlikely to be sufficient. Since medical technology changes over time, as old procedures are refined and new ones are developed, evaluating care for a particular condition necessitates continual updates on relevant technologies.

The dearth of such information on medical technologies is well known. OTA and others have previously documented the enormous gaps in knowledge about new and existing technologies and have developed relevant policies (53,103,452, 453,628). Although medical technology assessment deserves continuing attention and improvement, this report takes the deficiencies as given, but does not discuss them thoroughly or present policy options to address them directly.

Although the scope of this report is limited to quality assessment and does not extend to quality assurance, the two are closely related. *Quality assessment* measures and perhaps monitors the quality of medical care, while *quality assurance* seeks to safeguard and improve quality (186,384). Historically, much of the interest in assessing quality has come from concern about assuring quality, and many of the present activities related to quality fall under the rubric of quality assurance. Some of these, such as a hospital's procedures to screen the credentials of physicians for the staff, relate to the design of the system, while others, such as review of records by hospital committees and governmental bodies, are intended to monitor providers' performance and to take any corrective action required.

More generally, industries other than health care have developed a notion of quality improvement that entails companies' working with organizational and individual consumers to improve quality. The responsiveness of a company to consumers is an essential feature of quality control in these industrial programs and might be transferable to medical care delivery (68).

The results of quality assessment may feed into quality assurance and quality improvement through the responses of hospitals and physicians, employers, third-party payers, and Federal and State governments to problems that are identified. Indeed, some experts regard how a provider responds over time to deficiencies in quality as a measure of that provider's quality (67). In its evaluation of hospitals, the Joint Commission on the Accreditation of Healthcare Organizations (JCAHO) has examined how institutions have dealt with deficiencies in performance or other problems that have arisen. As part of its effort to develop clinical and organizational indicators of quality, JCAHO plans to monitor on a continuing basis how hospitals respond to recognized problems (329).

In this report, OTA assesses eight categories of potential indicators of the quality of care provided by physicians or hospitals (see table 1-1):

- hospital mortality rates, for the institution overall, by department, and by condition or procedure;
- adverse events that affect patients, as exemplified by nosocomial (institutionally acquired) infections in hospitals;
- formal disciplinary actions by State medical boards against physicians, sanctions imposed by the U.S. Department of Health and Human Services (HHS) on the recommendations of utilization and quality control peer review organizations (PROs), and malpractice compensation;
- evaluation of physicians' performance through their care for a particular condition, as exemplified by hypertension screening and management;
- volume of services in hospitals and performed by physicians;
- scope of hospital services, with particular reference to emergency services, cancer care, and neonatal intensive care units;
- physician specialization; and
- patients' assessments of their care.

Table 1-1.—Indicators of the Quality of Care Evaluated by OTA, by Type of Medical Provider

Physicians
Adverse events
Formal State disciplinary actions
PRO/HHS sanctions
Malpractice compensation
Evaluation of physicians' performance: hypertension
Volume of services
Physician specialization
Patients' assessments

Hospitals
Hospital mortality rates
Adverse events
PRO/HHS sanctions
Malpractice compensation
Volume of services
Scope of hospital services
Patients' assessments

SOURCE: Office of Technology Assessment, 1988.

This report does not offer a comprehensive evaluation of the many quality indicators that have been suggested or used (185). Although OTA attempted to select the most promising indicators, without evaluating others, one cannot conclude that the eight categories of indicators considered by OTA contain the best measures in terms of validity and feasibility for consumer use. OTA chose indicators to reflect the perspectives of consumers and the medical, research, and policy communities. High priority went to indicators concerning aspects of care that matter greatly to consumers, such as humaneness and communication of information, and decisions that consumers are likely to face, such as selecting a hospital to provide emergency care. To reflect policy interests, OTA paid particular attention to indicators that quality assessors are using or considering, especially for public programs. The indicators also illustrate different approaches to measuring quality and cover different aspects of quality. To ensure the feasibility of its own analysis, OTA limited its choice to indicators for which sufficient information existed to support an evaluation.

The remainder of this chapter summarizes the body of the report and presents policy options to address the problems identified. The summary first discusses the audience for information on quality, the assessment of quality from an individual consumer's perspective, and the dissemination of information to individuals. The summary then turns to the findings and conclusions regarding the specific indicators evaluated in this report. Based on the issues raised in the summary, the final section of this chapter analyzes policy options for congressional consideration. The body of the report considers the dissemination of information on quality, develops a framework to assess quality for individual consumers, and evaluates the eight categories of indicators. Appendix A describes the method used to conduct the study, and appendix B acknowledges the valuable assistance of many individuals. Appendix C presents the method that OTA used to analyze the reliability, validity, and feasibility of the indicators evaluated in this report. Appendix D discusses quality assessment activities of the Joint Commission, the American Medical Association, and the PROs, and appendix E lists recent and ongoing research on quality assessment in selected public and private organizations.

SUMMARY

Many individuals and organizations that make decisions about the purchase and provision of medical services could use valid information about the quality of medical care to guide their choices. Individuals seeking medical care have historically relied on family and friends for advice and on physicians for referrals to other medical providers. All to one degree or another have lacked information on quality. Quality-of-care information is also important for employers and third-party payers who monitor the performance of physicians and hospitals or who selectively contract with certain providers. Unions would like to have information on quality to evaluate alternative plans and providers for their members, especially as cost-containment efforts have led insurers and employers to increase cost-sharing and to favor lower cost providers. With information on quality, these organizations could consider quality as well as costs in their selection of and arrangements with providers.

Physicians, hospitals, and other providers themselves have lacked information on the quality of care and could benefit along with consumers from improved sources of information. Physicians could use valid information on the quality of care to select hospitals for staff appointment and for patient referral, to select other physicians for patient referrals, and to answer questions and interpret data for patients. Hospitals could also benefit from improved information about quality, in appointing physicians to staff and granting physicians admitting privileges and in monitoring their own performance and augmenting their quality assurance and risk-management programs.

Quality From the Perspective of Individual Consumers

Although many purchasers and decisionmakers can use information on quality, medical care is intended to benefit individual consumers. Thus, it is appropriate to evaluate the quality of care from the perspective of those individuals.

The quality of medical care has many dimensions, a fact that reflects the diversity of acceptable outcomes for patients and the complexity of the medical care process. Medical care seeks to promote, maintain, and restore people's health (186). Health itself contains multiple dimensions, including physiologic health, physical functioning, mental health, and social functioning. Depending on their conditions, patients vary widely in the health outcomes that they desire, from increased longevity, mobility, and emotional wellbeing to reduced illness, deterioration, and suffering. The appropriate content of care varies accordingly, from prevention and screening to diagnosis, rehabilitation, counseling, and other therapy. Moreover, patients vary in their preferences; some prefer less-invasive, less-painful, or less-disfiguring technologies, even at the expense of a shorter life.

Reflecting the complexity of the medical care process, prominent scholars have stressed the importance of evaluating both technical and interpersonal aspects of care (105,183). Both technical care, the application of medical science and technology to a problem, and interpersonal care, the personal interaction between patient and provider, enter into any episode of care and merit evaluation. Although consumers, providers, and the overall society from their own perspectives may emphasize different aspects of quality, all view both the technical and interpersonal aspects as important (183). Physicians have usually confined their evaluation to technical performance, while patients have shown more sensitivity to how they are treated (186). Society has more interest than individual consumers or providers in the equitable distribution and public health benefits of care, such as prevention of communicable disease.

Besides encompassing the many dimensions of medical care and health outcomes, a definition of quality must take into account the limits and continuing evolution of medical knowledge. As knowledge expands, some technologies, such as gastric freezing to treat stomach ulcers, become obsolete and should be discarded, while others, such as cimetidine, are shown to be efficacious and should be adopted as appropriate therapies. The use of medical technology also entails some risk and cannot guarantee improvement in a patient's health. In a larger sense, the uncertainty surrounding patient outcomes stems from the fact that medical care is but one influence on the health of an individual or a population. In fact, an individual's genetic makeup, environment, and lifestyle seem to play a greater role than medical care in explaining the causes of death and illness that now predominate in the United States.

The triad commonly used to assess the quality of care focuses on the structure, process, and outcome of care (183). Table 1-2 categorizes the indicators evaluated by OTA according to the assessment approach.

Table 1-2.—Indicators of the Quality of Care Evaluated by OTA, by Assessment Approach

Structure
Volume of services
Scope of hospital services
Physician specialization
Patients' assessments
Process
Adverse events
Formal State disciplinary actions
PRO/HHS sanctions
Malpractice compensation
Evaluation of physicians' performance: hypertension
Patients' assessments
Outcomes
Hospital mortality rates
Adverse events
Formal State disciplinary actions
PRO/HHS sanctions
Malpractice compensation
Evaluation of physicians' performance: hypertension
Patients' assessments

SOURCE: Office of Technology Assessment, 1988.

The *structure of care* encompasses the resources and organizational arrangements in place to deliver care, such as medical personnel, facilities, and quality review committees. Assessing quality via structural indicators, such as physician specialization, presupposes that their presence increases the likelihood that providers will perform well and their absence, the likelihood that providers will perform poorly. This assumption in turn raises the question of whether specific structural characteristics are, in fact, associated with better performance.

The *process of care* refers to the activities of physicians and other health professionals engaged in providing medical care. Although the appropriate care for a specific condition changes as knowledge expands, the thorniest problem with process measures of quality lies in the paucity of information about the efficacy of even well-accepted medical procedures. One should limit evaluations of providers' performance to procedures likely to improve or harm patients' health and satisfaction. The problem is that the link between the process of care and patient outcomes has been established for relatively few procedures.

Measuring quality via *outcomes*, namely changes in patients' satisfaction and health status, is the third approach. The problem with this method is that attributing changes in outcomes to medical care requires distinguishing the effects of medical care from the effects of the many other factors that influence patient health and satisfaction.

In light of the conceptual difficulties just mentioned, process and outcome measures should be regarded as complements rather than alternatives to assess quality. Process measures gain validity as quality indicators only to the extent that they have been found likely to improve patient outcomes, and outcome measures gain validity only to the extent that they have been linked to the prior medical care process. Similarly, to acquire validity as indicators of quality, structural measures must be shown to be associated with efficacious medical processes or validated outcomes.

Over the years, scholars have taken many different approaches to incorporating these complexities into a definition of the quality of medical care. This report examines several possible indicators of the quality of care provided by hospitals and physicians. Reflecting this task and the points above, this report uses the following definition of quality to guide its discussion: The quality of medical care is the degree to which the process of care increases the probability of outcomes desired by patients and reduces the probability of undesired outcomes, given the state of medical knowledge.

Under this definition, the quality of a hospital's or physician's care is judged against the likelihood that the care will achieve the desired patient outcomes. Which elements of patient outcomes (health and satisfaction) predominate depends on the individual patient or condition. As emphasized above, valid assessments of quality require linking the medical care provided (the process of care) with the effects on patient health and satisfaction (the outcomes of care).

This definition of quality also incorporates the notion that there are different levels of quality: a minimum level below which quality is unacceptable and levels of acceptable quality, including some levels in which important concerns about the quality of care remain and improvement is possible. Quality assessment and information systems take on different purposes that correspond to the different levels of quality: to identify unacceptable providers, so that they can be helped to improve and, as a last resort, be removed from practice; and to identify gradations among good quality providers, so that people can gravitate to the better ones and perhaps ultimately improve the general level of care. Since consumers vary in the importance that they attach to different aspects of care, information systems could also identify discretionary aspects of practice, so that people could act on their preferences.

A framework to assess quality from a consumer perspective starts with the technical and interpersonal aspects of care that influence desired outcomes, namely improvements in the various dimensions of health and in patient satisfaction. Such a framework should also address the choices that people face and the care that they receive during an episode of care. Surveys of individual consumers and the literature indicate that the following aspects of the medical care spectrum have importance for patient health and satisfaction:

- responsiveness to urgent and emergency situations;
- referral to the appropriate level of care;
- humaneness;
- communication of information;
- coordination and continuity of care among providers;
- primary prevention;
- case finding;
- evaluation of the presenting complaint;
- diagnosis; and
- management of the condition, which may include patient education, referral and consultation, therapy, monitoring, and followup.

Photo credit: American College of Emergency Physicians

Providers' responsiveness to emergencies and providers' referral of patients to the appropriate level of care have strong implications for patients' outcomes.

Although this report generally excludes issues of access, two aspects of access clearly overlap with quality of patient care and have such strong implications for patient outcomes that they are included in this report: providers' responsiveness to urgent or emergency care and providers' referral of patients to the appropriate level of care. Inclusion of the next two aspects in the framework reflects that people place high priority on being treated with respect and on receiving pertinent information from their physicians, including information to prevent disease and promote health (392). The last five categories, from primary prevention to management, relate to steps in the medical care process. Coordination of care receives separate mention to emphasize that, even if each step in the process is performed appropriately, poor-quality care can result from lack of coordination among providers. Continuity of care improves patient satisfaction and compliance (177), although its importance, like that of other aspects of care, varies with the situation (183). The relationship between these aspects of care and the indicators evaluated by OTA is summarized in table 1-3.

There is limited evidence on how quality-of-care information is likely to affect people's choice of providers. No empirical study addresses directly the effects of such information on consumers' choices or the elements of an effective strategy for disseminating such information. But drawing on principles of health behavior and studies in related fields, one may hypothesize that the following elements are necessary for consumers to receive information and to incorporate it into their choices of physicians and hospitals:

- stimulate consumer interest in the quality of care,
- provide information easy to comprehend,
- use many media and formats to present the information,
- use respected sources to interpret the information,
- make the information readily accessible, and
- provide consumers the skills to use and physicians the skills to provide the information.

These elements, like the studies from which they were drawn, relate mainly to mass communication. Although mass media have a role to play in raising consumers' awareness of quality-of-care issues and information, approaches that also included social support and skills training are likely to prove more effective in stimulating people to apply quality-of-care information to their interactions with providers and their choices regarding particular medical problems.

Findings Regarding Specific Indicators of Quality

Although none of the indicators evaluated in this report convey definitive information about the quality of an individual hospital or physician across the range of medical care, several of these indicators can provide useful information to organizations and individuals. For those consumers who consider physicians' character as well as skills in judging the quality of care, formal disciplinary actions by State medical boards can be accepted as valid indicators of poor-quality physicians. Consumers and others would be well advised to use many of the other indicators as initial screens for possible quality problems and to combine information from several indicators to decide whether further exploration is warranted. Information about unacceptable care merits more attention than information that ranks good-quality providers because of the more immediate concerns raised by poor quality and the state of quality assessment techniques.

Used as screens, certain indicators can identify physicians or hospitals about which there are reasonable grounds for concern. Armed with this information, individuals could then question their providers and evaluate whether a quality problem exists. A hospital whose unadjusted mortality rate exceeds expected levels, for example, may house a regional trauma center; this factor rather than poor quality might account for the high mortality rate. Similarly, that a hospital recommended by a surgeon has a low volume of cardiac surgery may reflect accounting conventions and not be related to the quality of care.

Consumers would also be well advised to combine information from more than one indicator of quality, to increase the likelihood of learning whether a quality problem was or was not present. A cardiac surgery patient could gain confi-

Table 1-3.—Indicators of the Quality of Care Evaluated by OTA, by Aspect of Medical Care

	Overall performance	Responsiveness to urgent situation	Referral to appropriate level of care	Humaneness	Communication of information	Coordination and continuity of care	Primary prevention	Case finding	Evaluation of presenting complaint	Diagnosis	Management
Hospital mortality rates	X	X	X			X		X	X	X	X
Adverse events	X	X	X		X	X	X	X	X	X	X
Formal State disciplinary actions	X	X							X	X	X
PRO/HHS sanctions	X	X				X		X	X	X	X
Malpractice compensation	X	X	X	X	X	X		X	X	X	X
Evaluation of physicians' performance: hypertension			X	X	X	X	X	X	X	X	X
Volume of services	X					X					X
Scope of hospital services		X	X			X					
Physician specialization	X										
Patients' assessments	X	X		X	X	X	X		X	X	X

SOURCE: Office of Technology Assessment, 1988

dence if a hospital performed a substantial number of relevant procedures, the hospital had a low mortality rate, and the surgeon had extensive training and experience in the procedure. By the same token, if the hospital had a high mortality rate and a low volume of procedures, the patient might wish to question the surgeon about that hospital and about alternatives, even if other hospitals required longer travel.

This approach poses certain difficulties, however. Patients may be reluctant or lack the skills necessary to raise such questions with their physicians. Furthermore, physicians may not be a reliable source of information about their own quality and may not have the knowledge to interpret information about other providers. Consequently, publicizing information on quality indicators may erode patients' trust in their providers, perhaps unduly if no quality problem exists. The response of providers, organizational purchasers, and consumer advocacy groups to information on quality may prove more productive. Inquiries by organized purchasers, such as employers and third-party payers, and consumer advocates would most likely spur providers to examine their performance. These groups may have medical experts and methodologists to interpret the information and have more leverage to exert through their market share. If indicators suggested problems with the hospitals or physicians to whom physicians referred patients, physicians could explore the situation and might decide to change their referral patterns. One would hope that hospitals and physicians about whom the indicators raised concern would examine their own practices and resolve any quality problems detected.

Table 1-4 summarizes the key findings regarding the indicators evaluated in this report. Hospital mortality rates and the adverse event nosocomial infections in hospitals show promise as indicators of quality. Up to one-third of hospital deaths and nosocomial infections in hospitals may be preventable (190,272). These findings emphasize the importance of a two-step process: first, to collect data about an adverse event and second, to examine medical records to determine whether a quality problem exists. Quality assessment techniques have not progressed to the point that one may rely on outcome data alone. For example, one analysis of medical records in hospitals with above average mortality rates identified quality deficiencies in 3 percent of all cases (462), and another analysis detected fewer problems in high-mortality hospitals than in other hospitals (279). One study that reviewed medical records and adjusted for patients' risk of dying did find that high-mortality hospitals were significantly more likely to have quality problems than low-mortality hospitals (190). Although researchers have identified characteristics of patients at high risk of dying in intensive care units and of contracting nosocomial infections in hospitals, techniques to adjust for patient risk across the hospital for all conditions are still being developed and tested. Furthermore, the generic quality screens that PROs use to review Medicare cases for the Health Care Financing Administration (HCFA) have not been validated.

The rigorous due process followed by State medical boards lends credibility to the validity of their formal disciplinary actions against physicians. State boards are reluctant to censure physicians and accord accused physicians extensive opportunity for appeal. In reviewing the cases of Medicare beneficiaries, PROs also follow a rigorous process, although it is newer and still undergoing refinement. For both formal disciplinary actions by State medical boards and PRO/HHS sanctions, the grounds for censure go beyond incompetence and inappropriate care to include felony, fraud, and impairment from drug abuse for the former and improper documentation for the latter. Opinions vary about whether these additional grounds relate to the quality of care.

Single incidents of malpractice compensation have little significance for a provider's technical quality, but repeated awards, especially for similar errors, justify attention. A malpractice suit more clearly indicates a patient's dissatisfaction with a provider's care, especially interpersonal aspects. But besides a provider's negligence, many factors related to judicial and insurance procedures determine the outcome of a malpractice suit, and in some specialties, such as obstetrics-gynecology, the vast majority of physicians have been sued. Physicians' malpractice profiles should be considered by specialty to take into account, albeit in a crude way, the fact that procedures and specialties vary greatly in the risk posed to patients, the likelihood of poor patient outcomes and malpractice suits, and the difference in malpractice compensation across specialties. Even if a rare event, such as a malpractice award, were distributed randomly among physicians, on statistical grounds one would expect a small number of physicians to account for a substantial number of cases. Furthermore, data on factors, such as physicians' caseloads and other characteristics, that could influence malpractice rates independently of the quality of care, are insufficient to permit attributing differences in malpractice rates to differences in the quality of physicians' care. As is the case for other indicators of quality, considering physicians' malpractice profiles over several years may dampen the influence of extraneous factors and reveal patterns more indicative of technical quality. People may also gain greater insight by combining information about malpractice compensation with other related indicators, such as adverse events and disciplinary actions.

Evaluations of a physician's performance for a specific condition can produce valid assessments of quality, if, as is the case with hypertension screening and management, the assessment criteria have been linked to changes in immediate outcomes, such as physiologic effects, or in more long-run aspects of patients' health and satisfaction. The most reasonable approach to evaluating medical care provided by physicians is a combination approach using both explicit criteria and the implicit judgments of experts to review the process of care and perhaps using patient outcomes to target the cases selected for review. This combination approach has not been well evaluated. Furthermore, an evaluation of a physician's performance for one condition is not necessarily generalizable to the physician's other conditions or to the physician's overall practice. Efforts underway in the United States and Canada to evaluate physicians' performance across a range of conditions are promising.

Researchers have examined whether the volume of services in a hospital or performed by a physician is associated with differences in patient outcomes, such as mortality. For certain procedures, such as coronary artery bypass surgery and total hip replacement, researchers have found lower volumes in hospitals to be associated with higher inhospital mortality rates or other adverse patient outcomes. By contrast, researchers have not documented this relationship for the volume of services performed by physicians or for all the services studied. Nor has the association between lower volumes and worse outcomes been validated by linking lower volume to deficiencies in medical care. Because the relationship is between volume and patient outcome, adjusting data for patients' risk poses the same problems here as for hospital mortality rates. As with several of the other indicators, consumers and others would be well advised to consider hospital volume data for more than a single year and to consider volume along with other indicators, especially the mortality rates of specific hospitals. Low mortality rates for cardiac surgery in a hospital with low volumes, for example, would be reassuring, in contrast to a pattern of high mortality rates and low volume.

External standards and guidelines based on expert opinion appear to provide a reasonable basis for assessing the adequacy of a hospital's scope of services, such as emergency rooms or neonatal intensive care units. Although a hospital's compliance with external standards or guidelines for scope of services has not been validated as a quality indicator through process or outcome measures, it seems worthwhile for consumers to seek hospitals judged by independent experts to have the appropriate resources to provide care, either overall or for specific conditions.

Although certification by a medical specialty board has not been associated with the quality of a physician's care, physicians practicing in the area of their training are likely to deliver higher quality care. Recertification of physicians over time, expanding the certification process to evaluate clinical competence, and limiting the designation of specialist to physicians with certain training and experience would improve the validity of physician specialization as an indicator of quality.

Patients' ratings provide valid information about the interpersonal aspects of and patients' satisfaction with physicians' ambulatory care and physicians' and hospitals' inpatient care. Although

Table 1-4.—Summary of Key Findings on Quality-of-Care Indicators Evaluated by OTA

Indicator	Strengths	Weaknesses
Hospital mortality rates	• A substantial percentage of hospital deaths are preventable. • There has been some limited validation of association between high mortality rates and poor performance. • Regulations address some undesirable provider behavior encouraged by use of this indicator. • Limited information is now publicly available.	• Techniques to adjust for patients' risk are inadequate. • Clinical data to adjust for patients' risk are not readily available. • Using this indicator to measure quality may result in many false negatives and false positives. • Diagnostic data are not uniformly coded and collected. • Many lay and medical people lack sufficient knowledge to interpret data on hospital mortality.
Adverse events, including nosocomial (hospital-acquired) infections	• A substantial percentage of adverse events are preventable. • Nosocomial infections have been partially validated as indicators of quality, and the characteristics of patients at high risk of nosocomial infections have been identified. • Infections of surgical wounds can be measured more reliably than all nosocomial infections. • Two-stage systems of screening for adverse events and auditing medical records are already in widespread use in hospitals; the cost of implementing the use of adverse events as quality indicators would be low. • Data from PROs' applying HCFA's generic screens are regularly compiled.	• Case finding of nosocomial infections is unreliable across hospitals. • No two-stage system, including HCFA's generic screens used by PROs, has been completely validated for evaluating quality across hospitals. • Using adverse events as quality indicators results in many false positives. • Screening and incident reporting vary considerably in the criteria used to identify adverse events; data collection and reporting are not uniform. • If use of adverse events as an indicator depends on providers' reporting adverse events, there is a high potential for gaming.
Formal disciplinary actions by State medical boards against physicians	• The indicator gains credibility from the rigorous due process used by State medical boards. • Grounds for disciplinary actions extend beyond incompetence and inappropriate care to felony, fraud, and impairment from drug abuse; if one accepts these grounds as relevant to quality, formal disciplinary actions are valid indicators of poor-quality physicians. • Information on formal disciplinary actions is already available.	• The precision of the grounds for disciplinary actions varies among State medical boards. • The indicator does not identify all poor-quality physicians; there are many false negatives. • Information on formal disciplinary actions is not well publicized in most States.
PRO/HHS sanctions	• The PRO/HHS sanctioning process is a rigorous one. • Most grounds for sanctions relate to incompetence and inappropriate care.	• The PRO/HHS sanctioning process is new, evolving, and sometimes unclear; information about grounds for sanctions is not easily accessible. • New methods of disseminating information on PRO/HHS sanctions have not been evaluated. • The grounds for sanctions may relate to improper documentation by providers, which some may not deem to be related to the quality of care.
Malpractice compensation	• Malpractice compensation indicates patient dissatisfaction. • Multiple jury awards justify attention. • Information on jury awards exists.	• Single incidents of malpractice give little indication of the technical quality of care. • Malpractice compensation is not a very reliable measure of quality. • Many factors unrelated to merits of a malpractice case affect its outcome. • Data are not available to adjust the results of malpractice cases for factors other than poor-quality care that may influence the outcomes. • Information on malpractice events is not routinely compiled and publicized.
Evaluation of physicians' performance for a specific condition, such as hypertension, by process or outcome measures	• Evaluations that combine explicit criteria and implicit judgment to review the medical care process, perhaps with patient outcomes to target review, hold promise as an indicator of quality, but are not well evaluated. • Evaluations across a range of medical conditions are promising, though not well evaluated. • Having expert panels develop criteria and standards for evaluating physicians' performance appears reasonable.	• Developing criteria and standards for evaluation requires prior proof of the procedures' efficacy; such proof is not available for many conditions. • The validity of criteria and standards developed by expert panels has not been evaluated. • The generalizability of the results of an evaluation to other settings and conditions is low. • Interpersonal aspects of care are not well represented in medical records and have not been well evaluated in reviews.

continued on next page

Table 1-4.—Summary of Key Findings on Quality-of-Care Indicators Evaluated by OTA—(cont'd)

Indicator	Strengths	Weaknesses
(continued from above)		
Evaluation of physicians' performance for a specific condition, such as hypertension, by process or outcome measures		• The explicit portion of the review of physicians' performance raises the problem of false negatives. • The implicit portion of the review of physicians' performance raises the problem of the reliability of physicians' judgments. • Data in patients' charts are not uniformly recorded; data on insurance claims are not uniformly coded, collected, or reported. • Publicizing results of peer review may impair physicians' willingness to participate in the review and to be candid.
Volume of services in hospitals or performed by physicians	• Lower hospital volumes have been associated with higher rates of poor patient outcomes for certain services, mostly surgical. • Data on hospital volume are readily available from claims or hospital discharge abstracts; extra cost of data collection would be low.	• Data on diagnoses and other patient characteristics are not uniformly coded, collected, or reported. • A relationship between lower volumes and higher rates of poor patient outcomes has not been documented for services performed by physicians or for all services in hospitals. • A relationship between volume and outcome has not been validated by linking lower volume to poor-quality care. • It is not clear that patient differences have been adequately taken into account in studies of the volume-outcome relationship. • Using volume of procedures as an indicator of quality would give providers an incentive to raise volume by relaxing standards of use.
External standards and guidelines for scope of hospital services, including emergency rooms, cancer care, and neonatal intensive care units	• Standards and guidelines developed by external experts are a reasonable means for assessing minimum acceptable resources to manage certain conditions. • Some information on the indicator is collected and publicly available.	• The use of the indicator to measure quality has not been validated through process or outcome measures. • Information on the indicator is difficult to obtain.
Physician specialization as measured by specialty board certification or by practicing in one's area of training	• Practicing in one's area of training has good validity as an indicator of the quality of technical aspects of care. • Information on the training of broard-certified physicians is readily available. • Requiring periodic recertification and expanding certification to include clinical competence are promising methods to improve the validity of board certification as an indicator of quality.	• The association between practicing in one's area of training and providing better quality care is not generalizable to other specialties, diagnoses, or procedures. • The relationship between practicing in one's area of training and interpersonal aspects of quality has not been studied. • Information on the training of non-board-certified physicians is not readily available. • Board certification is not a valid measure of quality. • With the use of physician specialization as an indicator of quality, the potential for gaming is high if physicians may designate themselves specialists.
Patients' assessments of their care	• Patients' ratings are a valid indicator of the quality of interpersonal aspects of care and of patients' satisfaction with physicians' ambulatory care and physicians' and hospitals' inpatient care. • Patients' assessments relate to good and poor care and to access. • Patients' ratings and reports of technical aspects of care are promising as quality indicators, especially for physicians' ambulatory care, but they have not been validated.	• Adequate data collection methods and instruments have not been developed and standardized. • Potential bias in assessments may result from patients' preferences or other characteristics. • Special surveys are required to collect data.

SOURCE: Office of Technology Assessment, 1988.

Photo credit: Harvard Community Health Plan

Patients' ratings provide valid information about interpersonal aspects of physician and hospital care and are promising indicators of technical aspects of physicians' ambulatory care

less information exists about patients' ratings and reports of technical aspects of care, they appear promising, especially for physicians' ambulatory care. Patients' assessments relate to both positive and negative aspects of care and can provide information about access. Like other outcome measures, however, patients' ratings may reflect factors other than quality, such as the preferences of the particular patients in a physician's practice.

Although not thoroughly validated in this report, certain situations suggest quite strongly that hospitals or physicians are providing care well below minimum acceptable levels of quality. The Joint Commission on the Accreditation of Healthcare Organizations (JCAHO) refuses to accredit 1 to 2 percent of the hospitals that it surveys (524). Ninety percent of the hospitals surveyed by JCAHO are accredited with contingencies relating to deficiencies that the hospital is to correct within a certain period of time. Since JCAHO refuses accreditation only to hospitals with substantial failings, such refusal may be taken as an indication of a poor-quality hospital. Hospitals or offices in extreme disrepair, perhaps as an outgrowth of financial difficulties, also suggest poor quality. More specific to a particular condition, hospitals that have high birthweight-specific mortality rates probably offer lower quality care for newborns than hospitals with lower rates. Physicians who continue to perform outmoded procedures, such as those on the list developed by the National Blue Cross-Blue Shield Association, or physicians who perform complex surgery or other complex procedures without appropriate training and experience are likely to offer care of low quality.

In the course of evaluating specific quality indicators, this review has identified several deficiencies that pervade the field of quality assessment. Current techniques cannot adequately adjust for patient and environmental factors that may influence patient outcomes independently of the quality of care. This situation greatly impedes the use of outcome measures, such as hospital mortality rates, as indicators of quality. Nor has research validated possible quality indicators by linking structural measures of quality with appropriate process and desired patient outcomes, process measures of quality with subsequent patient outcomes, or desired patient outcomes with prior process. Although this report attempted to develop a framework for assessing quality from an individual consumer's perspective, a conceptual framework is still lacking for the most likely hazards of medical care, to indicate how medical care is likely to fail and how to test for each major failure.

Intertwined with the shortcomings of assessment techniques is the dearth of necessary data to assess the quality of care. Several of the indicators—hospital mortality rates, adverse events, malpractice compensation, evaluation of physicians' performance through a specific condition, volume of services, physician specialization, and patients' assessments—suffer from lack of uniform methods to code, collect, and report data, especially about specific diagnoses. No routinely collected data permit quality assessors to evaluate physicians' practices outside of hospitals. Additional data (such as diagnostic information for Medicare ambulatory care) and methods (such as uniform reporting requirements) are needed to assess the quality of ambulatory care. Although more information is available on hospital than ambulatory care, the hospital discharge information required by Medicare contains little information on the patient's status on admission, before the person received care. Even with Medicare data sets, one cannot easily track the services that a patient has received from different providers on an inpatient and ambulatory basis. This deficiency makes it difficult to attribute specific patient outcomes to prior medical care, a problem that will intensify as care moves increasingly into ambulatory settings.

Although some information related to quality, most notably hospital mortality rates for Medicare patients, is becoming available to the public, other relevant information, such as the JCAHO contingencies that a hospital receives, is regularly compiled but not publicly available. Nor is information covering several years generally available on certain quality indicators, such as hospital mortality rates, adverse events, and volume of hospital procedures. Such longitudinal information would be less likely than information for a single year to reflect random influences and more likely to indicate relationships related to the quality of providers' care. Some current efforts are beginning to periodically generate information, such as the hospital mortality rates of Medicare beneficiaries, and systems could be established to regularly produce information on other indicators.

When considering making quality-of-care information more generally available, one must consider the likely effect on medical providers. The use of some indicators may create perverse incentives. In the absence of techniques that adequately adjust for patient differences, for example, evaluating the quality of hospitals by their mortality rates would entail an incentive for hospitals to transfer or avoid admitting severely ill patients. Similarly, using hospital-acquired infections or other adverse events as indicators of quality could undercut efforts to diagnose, document, and correct certain deficiencies. The same effect could arise from applying criteria to evaluate physicians' performance for a specific condition, such as hypertension. Evaluating hospitals or physicians by the volume of procedures that they perform might encourage them to relax their criteria for using these procedures and perhaps perform some unnecessarily. These are but a few examples of how a conflict might arise between a climate to encourage hospitals and physicians to examine and improve their care and efforts to make assessments of providers' quality more publicly available. In some cases, regulations have addressed a problem, such as hospitals' transfer of severely ill patients, but such regulations have not resolved the underlying conflict. This conflict is particularly troubling because most reviews of medical care, both public and private, rely on physicians and other medical professionals and will continue to do so.

POLICY OPTIONS

This report has identified some potentially useful indicators of the quality of care, but also several deficiencies associated with quality assessments to guide consumers' choice of hospitals and physicians. The remainder of this chapter examines approaches that Congress could take to remedy the problems noted above in five areas: to improve techniques available to assess the quality of care, to ensure that acceptable techniques are used to produce quality assessments, to improve the availability of data required for quality assessments, to disclose information to the public, and to disseminate information on quality to individuals and organizations. Policy options in each of these areas are summarized in table 1-5. These approaches represent policy options, not recommendations, for Congress. Although some of the options are related, others are mutually exclusive approaches to address a particular problem.

Table 1-5.—Summary of Policy Options for Congress To Address Problems With Quality-of-Care Indicators

To improve quality assessment techniques

Option 1: Mandate and fund research and demonstrations to improve quality assessment techniques.

To ensure the quality of quality assessments

Option 2: Mandate the selection of indicators to assess quality for Medicare and Medicaid.

Option 3: With option 2, mandate the use of indicators to assess hospitals and physicians in Medicare and Medicaid.

Option 4: With option 2, mandate briefings of State and local groups on selected indicators and construction methods.

To improve the availability of required data

Option 5: With options 1 and 2, require demonstrations to collect clinical data from hospitals and physicians to assess the quality of their care.

Option 6: With options 1 and 2, establish a task force to develop uniform requirements for reporting data.

To disclose information to the public

Option 7: Require Medicare and Medicaid hospitals to make certain indicators public, including contingencies from JCAHO and results of HCFA's reviews.

Option 8: Permit PROs and HCFA to disclose information that identifies specific physicians.

To disseminate information to the public

Option 9: Establish an HHS office to disseminate quality information.

Option 10: Mandate and fund research and demonstrations on disseminating quality information.

SOURCE: Office of Technology Assessment, 1988.

Policy options must be considered in light of the fact that information on some of the indicators evaluated in this report is already being disseminated and used, namely information on hospital mortality rates, sanctions imposed by HHS on the recommendations of PROs, and physician specialization. As policymakers address problems of quality assessment, activities that will improve these indicators merit high priority, so that consumers and providers using current information are not misled. Moreover, efforts to identify and improve physicians and hospitals whose quality falls below acceptable levels deserve priority over efforts to distinguish among good-quality providers. Identifying poor-quality providers is not only more pressing for consumers and other providers, but also consistent with the obligation of the government to protect public health and safety and with the current state of quality assessment.

As the policy options illustrate, Congress could take three approaches, separately or together, to address these problem areas. One approach would be for Congress to create and maintain a legal climate conducive to the flow of information needed to evaluate providers' quality and to inform consumers. This approach would entail removing any legal barriers to providers' participation in quality assessment and to public disclosure of information useful to consumers. As a second approach, Congress could use the leverage of the Medicare and Medicaid programs to encourage hospitals, physicians, and States to undertake desired actions, such as collecting data, constructing indicators of quality, and making information publicly available. As a third approach, Congress could mandate that the Federal Government directly undertake efforts to remedy deficiencies regarding quality assessments for consumers.

Although whether a particular governmental activity is considered appropriate may depend on one's philosophy of government, consensus, if not unanimity, supports a government role in the flow of information. Scholars have often cited information to exemplify a good that is in everyone's interest to have but in no one's interest to finance individually. Like the responsibility for promoting public health and preserving national security, the responsibility for ensuring adequate vital public information may fall to government. This situation need not imply that the government itself undertake the desired activities. Some private sector organizations, notably the Joint Commission and the Institute of Medicine, already have considerable expertise and work underway. The Federal Government could stimulate private sector and State initiatives, promote the coordination of public and private activities, and cooperate in public-private enterprises. The discussion of the policy options below considers how Congress could encourage or use such non-Federal organizations.

Two relevant issues then arise for public policy: whether public information about hospital and physician quality has sufficient importance to justify governmental action and which approaches or options are likely to prove most effective in bringing about the desired results. As described earlier, individuals and organizations from many quarters support increased publicly available information on the quality of medical care for several reasons: so that consumers and providers can identify poor-quality physicians and hospitals, so that people can learn over time how to choose and interact with providers, and so that consumers through their choices over time can influence providers to improve their quality of care. The relative merits of different strategies to accomplish these ends are discussed under each option.

To Improve Quality Assessment Techniques

Although considerable work has been done to develop techniques to assess the quality of medical care, in general indicators require much refinement. The evaluation of the indicators in this report has brought to light several critical areas in which quality assessment techniques remain wanting. The inadequacy of techniques for taking into account factors other than quality that affect patients' outcomes impedes the public's interpretation of outcome measures, such as hospital mortality rates. Although the vast majority of medical care takes place in ambulatory settings, methods to assess physicians' ambulatory care are still in their infancy. Even more basic to quality assessment, the ability of structural and outcome indicators to measure the quality of care has not been validated by linking the results to the medical care process. Nor is there general agreement on the criteria and standards by which the medical care process should be judged. Identifying poor-quality providers is the immediate need, but techniques are also needed to distinguish levels of good-quality providers.

Option 1: Mandate and earmark funds for the Department of Health and Human Services, the Veterans Administration, and the Department of Defense to strengthen research and demonstrations to improve techniques for assessing the quality of medical care.

The Federal Government has a special interest in supporting quality assessment research. In addition to its role in developing basic research techniques, the Federal Government accounts for 30 percent of the Nation's medical expenditures, primarily through the Medicare program for elderly and disabled people and the Medicaid program for certain poor people, but also through the Veterans Administration for veterans and the Department of Defense for military personnel and their families.

Despite Federal and private funding of research on quality assessment (see app. E), serious gaps remain, and efforts do not flow from a systematic, long-term agenda. Few projects are attempting to validate outcome measures against the medical care that patients received or to examine the validity of structural measures of quality. Several projects are working on techniques to adjust outcome measures for relevant patient characteristics, but few of these plan to incorporate clinical information on a patient's status when the patient first sought medical care, information that is vital to assessing the quality of care that was subsequently provided. A continuing need to provide the basis for quality assessments of providers' performance is research on the clinical efficacy of common medical procedures. Currently funded projects do not appear to be laying the groundwork needed to assess the quality of medical care in ambulatory settings, an activity that the Omnibus Budget Reconciliation Act of 1986 (Public Law 99-509) stipulates that PROs are to undertake beginning no sooner than January 1989.

Yet another type of research needed to further the field of quality assessment is research on the criteria and standards for evaluating physicians'

Photo credit: March of Dimes Birth Defects Foundation

Information is lacking on the efficacy of many medical technologies, such as those used routinely in neonatal intensive care. Such information is needed to assess the quality of providers' performance.

and hospitals' performance. Drawing on the literature and expert opinion, some researchers have formulated criteria and standards to assess providers' performance for certain conditions. Generally accepted review criteria for many conditions are lacking, however, and quality assessors, including those in PROs, usually rely on their own criteria. For the most part, existing criteria and standards have not been tested. The generic screens that PROs apply to Medicare inpatient cases, for example, have not been validated. Nor has the process that PROs and the HHS Office of the Inspector General use to determine sanctions been evaluated. How to modify criteria over time to incorporate changes in medical knowledge and practice poses an additional challenge. Without some mechanism to take technological change into account, evaluating the quality of care through criteria and standards runs the risk of inhibiting medical advances.

Under the option described here, Congress not only would require the Federal agencies engaged in health services research and health care delivery to give high priority to research and demonstrations designed to improve quality assessment techniques, but also would earmark funds for this purpose. Federal agencies in turn could identify their research priorities and fund researchers to pursue them. Congress could rely on a decentralized research strategy, with each agency continuing to work independently. Alternatively, instead of continuing fragmented efforts in this field, Congress could establish a specific locus of responsibility for quality assessment research, in either an existing or newly created office.

That the Federal Government finances or provides medical care on a large scale gives it both the economic interest and the mechanisms to refine quality assessment techniques. The Government has considerable opportunity to amass data required for developing quality assessment techniques and to test alternative assessment methods across population subgroups, geographical regions, and medical care settings. Much could be learned by examining population-based data from Medicare. From these data, for example, researchers could derive statistics on the average and range of mortality rates for certain conditions. Those statistics could then be used to inform consumers of the risks of specific treatments and to serve as a benchmark for developing standards to evaluate providers. In addition to amassing useful data, the Government also has the ability to bring together experts from medical specialty societies and other parties at interest to develop criteria and standards for assessment.

This option raises the issue of what is to be gained from targeting funds or creating a new locus for quality assessment at this time. With the assistance of expert groups, the Joint Commission is developing and testing measures of clinical performance and patient risk (see app. D). The Omnibus Budget Reconciliation Act of 1986 (Public Law 99-509) mandated certain studies related to quality assessment, including one now underway at the Institute of Medicine to examine criteria and standards for assessment. Congress may prefer to await the results of these studies and others underway at HCFA and the National Center for Health Services Research and Health Care Technology Assessment (see app. E) before mandating that additional research on specific topics be undertaken. Alternatively, gaps in current research, such as work on survey instruments for patients' assessments of their care, could be identified and corresponding projects could be undertaken to avoid further delay.

To Ensure the Quality of Quality Assessments

Efforts to assess the quality of medical providers and to make the results public have mushroomed in recent years. In the Federal arena, the most notable effort was HCFA's release in 1986 and 1987 of the mortality rates experienced by Medicare beneficiaries in hospitals across the country. During those same years, the PRO for California disclosed publicly Medicare mortality rates in all California hospitals. Individual States, notably Massachusetts, New York, and Pennsylvania, are beginning to assess the quality of hospitals and physicians and plan to make the information public in the near future.

Private activities are also increasing. Individual hospitals, organizations of hospitals, large clinics, and HMOs are engaged in assessing the quality of their own care. These private organizations are using the results in part to identify their shortcomings and to improve the quality of their care. But some of them are also developing quality assessments for marketing purposes—to convince large employers or third-party payers to select their organizations as the preferred providers for certain procedures or for patients in certain localities.

Quality assessment techniques based on currently available data not only have many deficiencies but also are undergoing continual refinement. Thus, concerns arise about the technical skills of the individuals in both public and private organizations who are assessing quality. Do they have the requisite medical and statistical expertise? Are they able to hone their skills by incorporating new methods? For assessments performed by the medical providers themselves, additional questions arise about objectivity. By their very nature, those evaluations of a hospital that are developed for public relations or marketing are likely to promote that hospital and present biased information to the public.

The options discussed below would address these problems at several stages involved in assessing the quality of medical care: selecting indicators to assess quality, constructing those indicators, and upgrading the skills of the quality assessors. Options 2 and 4 involve having the Federal Government directly formulate or disseminate information on quality, while option 3 involves using the leverage of the Medicare and Medicaid programs and the example set by the Government to bring about desired changes.

Option 2: Mandate the Department of Health and Human Services, working with national experts, to select indicators for assessing the quality of care provided within the Medicare and Medicaid programs.

Selecting indicators for quality assessment requires technical expertise. As discussed in this report, indicators of quality vary widely in their reliability and validity and in the feasibility of their use. In the context of informing individual and organized purchasers about quality, selecting indicators to measure the dimensions of quality that matter to consumers assumes particular importance.

Under this option, Congress would require HHS to select indicators to assess quality in Medicare and Medicaid, the two Federal programs under its purview that finance or provide medical care. HHS would also be required to develop uniform methods of constructing the indicators selected. In all of these activities, HHS would draw on experts from quality assessment research and clinical medicine. Creating advisory groups of nongovernmental experts would improve the results and strengthen the credibility of the ultimate selections among the medical and quality assessment communities. Involving the medical community is particularly important to gain its support for using the indicators and methods chosen. Congress could extend the option to other Federal programs, such as the Veterans Administration and the Department of Defense. Government offices that assumed responsibilities under this option would require additional funding, because the work would necessitate additional staff and advisory panels.

Carrying out the requirements of this option would entail a continuing effort, to ensure that indicators were revised and updated as assessment techniques improved and new sources of data became available. HHS could limit the indicators initially to those currently in use, such as risk-adjusted hospital mortality rates and elements in the generic screens applied by PROs. HHS could then add indicators considered valid for which information already exists, such as contingencies that hospitals receive from the Joint Commission and disciplinary actions by State medical boards, and develop information on other valid indicators, such as patients' assessments of their care. Over time, as assessment methods advanced, HHS could revise methods of constructing indicators and add other indicators whose validity had been established or improved.

A disadvantage of this option is that the validity of many indicators of quality is dubious. A danger exists that HHS' decisions would entrench indicators and construction methods that measure quality poorly and lead to faulty evaluations of providers' performance.

Another drawback of this option is that Federal efforts in selecting indicators of quality and methods of constructing them could duplicate those already underway, especially at the Joint Commission and also at the Department of Defense. JCAHO is already developing clinical indicators in several areas and plans to expand to other clinical areas in the future. The Department of Defense currently uses explicit clinical criteria developed by panels of experts to review about 10 percent of all military hospital discharges. To avoid duplication of effort, HHS could include JCAHO and Department of Defense officials among its advisors and stay abreast of their evolving methods to evaluate quality. Alternatively, HHS could apply the JCAHO indicators, for example, as they are developed and tested, to evaluate care under the Medicare and Medicaid programs.

Option 3: If Congress has adopted option 2, mandate that the Health Care Financing Administration, using the indicators selected by the Department of Health and Human Services, annually assess the quality of the hospitals and physicians that participate in the Medicare and Medicaid programs.

This option would use the leverage of the Medicare and Medicaid programs to stimulate the use of the quality assessment indicators and methods selected under option 2. Congress could require that individual hospitals and physicians, as a condition of their being eligible to participate in Medicare and Medicaid, use standard assessment methods prescribed by HCFA and annually make the resulting information public; another approach would be for Congress to stipulate that PROs or States develop the information. Alternatively, HCFA could work with the Joint Commission, either by adopting the standards of clinical performance that are under development or by relying on the Joint Commission's accreditation process when it incorporates clinical indicators.

As discussed above, HCFA is already evaluating physician and hospital performance and making some information publicly available. This option would require that those efforts continue and make them part of a coordinated effort.

At least one State, Pennsylvania, has already taken action on informing the public about the quality of its health care providers. Under the Health Care Cost Containment Act of 1986 (Pennsylvania Act 1986-89), that State requires that statistics comparing hospitals and physicians on mortality, morbidity, infection, and readmission rates be published at least quarterly in generally circulated newspapers. Pennsylvania officials expect the first publication of some of the mandated data at the beginning of 1989 (82). Colorado requires that hospitals regularly submit data on patients' severity of illness and morbidity, and other States may follow. HHS could examine Pennsylvania's and Colorado's experiences as case studies to gain insights for its own programs.

Regardless of which entity is responsible for performing quality assessments, HCFA could specify uniform methods of data collection, techniques for adjusting data, and procedures for releasing and interpreting the results. If new tasks were added to the responsibilities of PROs or States, it would be vital for HCFA to train their staffs and for Congress and HCFA to increase their funding. If Congress were to require individual providers to develop and release the information, the quality of the information would probably be lower, but an increase in Federal funding might be avoided.

Requiring that information on the quality of care provided to Medicare and Medicaid patients be made available would yield information covering all age groups, from infancy and childhood through the childbearing years to old age. Moreover, Medicare and Medicaid patients probably represent the people at highest risk of developing complications from poor-quality care. People who are very old, disabled, or very young are least able to withstand physical insults. Poor people who are eligible for Medicaid and delay care for financial reasons may not obtain care until their medical conditions are fairly advanced. If the effects of poor-quality care are more likely to be manifested among Medicare and Medicaid patients, quality assessments based on these subgroups of the population will be especially likely to detect differences in the quality of medical providers. Some hospitals and physicians, however, treat few Medicaid patients. To broaden the range of patients covered, therefore, Congress may wish to stipulate that whoever constructs the indicators of quality incorporate data on all of a hospital's or physician's patients.

Like option 2, this option confronts the fact that quality assessment techniques are inadequate and many indicators have not been validated. If Congress does not wish to proceed with the indicators now being used and with those indicators that appear to give valid measurements of quality, it could reject this option and emphasize the improvement of assessment techniques, as outlined in option 1.

Option 4: If Congress has adopted option 2, require that the Department of Health and Human Services include in its State and local outreach activities briefings on the selected indicators and construction methods.

The requirements for assessing quality that option 3 would institute through the Medicare and Medicaid programs would not affect many ongoing activities to assess quality at the State and local level. Even now there exists a substantial gap between what researchers know about quality measurement and what employers and others think is available. In order to construct the indicators that HHS would recommend, the people undertaking quality assessments might well require additional training and skills. Nor can one expect that these technicians have access to information on refinements that are occurring in quality assessment techniques so that they can upgrade their skills.

This option would use the networks and skills that HHS has developed through other programs to disseminate information about quality assessment methods. Drawing on the expertise attained in the course of selecting and refining indicators to assess quality, HHS staff could improve the technical skills of quality assessors and, it is hoped, the results of their work. Since this option would add an area of responsibility to HHS activities, increased funding might be necessary.

Instead of relying on Federal Government staff to convey information and skills, HHS could work with State and private groups, perhaps through a clearinghouse. Business groups involved in quality assessment might be particularly interested in participating.

The problems of the inadequacy of quality assessment techniques and the paucity of validated indicators of quality arise for this option as well as others. In light of these problems, Congress and HHS may wish to delay outreach and training until better techniques have been developed and tested.

To Improve the Availability of Required Data

Attempts to assess quality flounder on more than the dearth of techniques. Interwoven with inadequate assessment methods is the inaccessibility of necessary data. Sometimes the requisite data exist, but not in a form readily available to researchers or quality assessors. Information on the admitting status of hospital patients presents a striking example. Judging the quality of hospital and physician care requires knowing a patient's condition when the person first sought care from a particular provider and the trajectory of that condition during a particular episode of care. Such information is a prerequisite to evaluating how a provider managed the condition and what role the quality of care played in the patient's eventual outcome. The only information routinely available, however, is information from hospital discharge abstracts, which report a patient's status and diagnosis *after* the patient has received care. Although minimum data sets have been de-

Photo credit: George Washington Medical Center

Data that are routinely available often lack the clinical details needed to assess the quality of care that patients have received.

veloped for ambulatory care, not even the equivalent of a discharge abstract is available for care provided by physicians in ambulatory settings.[1] Claims submitted to third parties for payment are promising sources of information for physicians' offices and other ambulatory sites, as for hospitals, but such claims often lack the clinical details needed to assess the quality of care.

Options 5 and 6 below present methods of addressing these problems. Option 5 involves conducting demonstrations through the Medicare and Medicaid programs to require providers to collect and report certain data. In light of the importance of coordinating any Federal requirements with those of States and private organizations, option 6 involves establishing a task force to develop uniform data requirements. Although the options could be undertaken independently, the demonstrations conducted under option 5 could test the requirements developed under option 6.

Option 5: Require the Department of Health and Human Services, as part of its research on quality assessment techniques in option 1 and its selection of indicators in option 2, to conduct demonstrations to collect from hospitals and physicians participating in Medicare and Medicaid whatever clinical data are needed to assess the quality of their care.

Although virtually all hospitals in the United States routinely create hospital discharge abstracts for each inpatient, routine access to this information has occurred only in the fairly recent past. In 1982, Medicare mandated that for all Medicare patients, hospitals participating in the program use Medicare's uniform bill (UB-82), which merges billing data with the standardized data elements and definitions agreed upon in the uniform hospital discharge data set. As of October 1987, 26 States had mandates to collect uniform information on all hospital discharges at the State level; 25 of these State data collection systems were established during the 1980s (366). Two additional States were collecting data for selected diagnoses.

In conjunction with efforts to improve quality assessment techniques under option 1 and to select indicators of quality in option 2, this option would require HHS through HCFA to conduct demonstrations through the Medicare and Medicaid programs to collect data required to evaluate certain techniques and to construct the indicators of quality being used. Hospitals, physicians, and other providers of acute care would be required to submit standardized information on each patient encounter. Identifying the data needed for quality assessment presupposes that HCFA has selected specific indicators and has ascertained which data are needed to construct them. Requiring that hospitals and other providers use uniform definitions for recording and reporting data would be vital to permitting subsequent comparisons across providers (see option 6).

Collecting and transmitting the data suggested in this option might entail sizable investments by providers in new data systems. Providers are already facing new data requirements from organizations outside the Federal Government. As part of its efforts to evaluate quality, for example, the State of Pennsylvania is requiring the hospitals in that State to report two data elements obtained from a specific software system to determine patient severity. The estimated cost to the average acute care facility for the first year of operation is $56,000, an estimate that does not include the routine costs of abstracting discharge data (82). Colorado requires collection of patient-level data on severity of illness at the time of admission. Although Colorado has not required hospitals to use a specific vendor's system, the State has stipulated that it eventually intends to collect uniform data (140). Iowa is actively pursuing a similar approach. JCAHO is contemplating changes in the data and data systems that it requires of hospitals. HCFA is already developing a uniform clinical data set for PROs to use on all cases reviewed (357). In order to minimize wasteful duplication, it would be essential to ensure that HCFA coordinated its requirements with those of the States, the Joint Commission, and others.

---

[1] Currently nine States have mandates to collect patient-level data from ambulatory care settings. Only Iowa and Maryland (both for hospital-based ambulatory surgery) are actually collecting data on patient encounters (using the Medicare uniform billing form). The other seven States have not yet implemented systems. None of the nine States will collect data from individual physicians' offices; instead, these States will collect information from ambulatory surgery centers, and sometimes from nursing homes or other freestanding ambulatory care centers, such as emergicenters (366).

Option 6: Require the Department of Health and Human Services, as part of its research on quality assessment techniques in option 1 and its selection of indicators in option 2, to establish a task force to develop uniform requirements for data to be reported by hospitals and physicians.

If Congress adopted option 1 to improve quality assessment techniques and option 2 to have HHS select indicators of quality, it would be important to foster the development of uniform requirements for reporting data that take into account the data needs of Federal agencies and other organizations. This option would take direct action to bring together the interested parties by creating a task force. The task force could be led by a private organization, such as the American Hospital Association, or by a governmental entity, such as the U.S. Committee on Vital and Health Statistics in HHS.

The task force would consist of experts in data collection, statistical analysis, and quality assessment plus representatives of hospitals and of organizations that routinely require or collect data, such as JCAHO, States, and third-party payers. The process of developing uniform requirements for reporting data could parallel the effort of the National Uniform Billing Committee that resulted in the creation and adoption of Medicare's uniform bill, UB-82.

Setting up a formal body to develop uniform requirements would increase the likelihood that the different organizations would adopt uniform or compatible data systems and that providers would accept the new requirements. The various organizations could also provide opportunities to test proposed changes before widespread implementation. The Federal cost of staffing a task force might exceed the cost of a separate HCFA activity, but the coordination achieved by the task force could well obviate substantial expenditures by medical providers and quality assessors, who would otherwise have to cope with different requirements and conflicting data systems.

A disadvantage of this option is that State and private organizations may be unlikely to adopt uniform methods of data collection and reporting unless required to do so. Even with activities during the 1970s, adoption of a uniform hospital discharge abstract is far from complete. If Congress considers eventual recommendations for uniform coding and reporting important, it could consider making them a condition that providers must fulfill to participate in the Medicare and Medicaid programs.

An additional drawback concerns the state of quality assessment techniques. Since methods are undergoing continual refinement, the Government could await the development of greater consensus on assessment techniques and their required data elements before attempting to reach agreement with other groups about uniform data requirements.

To Disclose Information to the Public

Some information relating to the quality of hospitals and physicians is routinely compiled but is not publicly disclosed. The options in this section would work through the Medicare and Medicaid programs to make the information available to the general public.

Option 7: Require as a condition of participation in Medicare and Medicaid that hospitals make publicly available information on certain indicators of quality, including the contingencies received from the Joint Commission on the Accreditation of Healthcare Organizations and the results of the Health Care Financing Administration's own review process.

In the course of their routine operations, hospitals develop information pertaining to quality. For example, hospitals can meet one of the Medicare and Medicaid conditions of participation through accreditation by JCAHO. Indeed, 78 percent of the hospitals paid under Medicare have taken that option (438). The other hospitals either failed to achieve JCAHO accreditation or chose the alternative procedure, to have HCFA through State agencies inspect and certify the institutions as having the necessary facilities and procedures to deliver acceptable care.

During the accreditation process, JCAHO applies criteria developed through consensus to evaluate aspects of almost all areas of a hospital (see app. D). The hospital receives contingencies for any areas that fall short of JCAHO's requirements. Although JCAHO refuses to accredit an institution with contingencies above a certain threshold, a hospital with fewer contingencies receives accreditation contingent on correcting the deficiencies within a period of time specified by JCAHO. Accredited hospitals receive a certificate for posting, and, upon request, JCAHO makes publicly available information on whether a hospital passes or fails to gain accreditation. JCAHO does not divulge, however, whether a hospital has any contingencies and, if so, the number of or reason for any contingencies (330). Individual hospitals may make that information available, but their policies vary.

Under this option, a hospital that chose to use JCAHO to satisfy conditions of participation would have to disclose any JCAHO contingencies, by number and area. HCFA or PROs could compile the information from all hospitals and annually release it, perhaps along with hospital mortality rates and any other similar information. Purchasers of health care, including individuals, third-party payers, and employers, could then question physicians and hospitals about the status of any contingencies, especially in areas related to procedures for which they were seeking care. Precedent supports public disclosure of this information. As part of the licensing process, 38 States require hospitals to give their State health agencies copies of JCAHO survey reports (48). At least one of these States, New York, will release copies of JCAHO reports to the public. The Veterans Administration also has a policy of making information about contingencies and accreditation available on request (146).

Photo credit: George Washington Medical Center

Physicians who practice in the area of their training are likely to deliver higher quality care than physicians without special training in the area.

Under this option, HCFA could require other quality-related information to be released, such as mortality rates for all the hospital's patients, HHS sanctions recommended by PROs, disciplinary actions by State medical boards, the total volume of certain services, and the training and experience of physicians performing specialized procedures. PROs or HCFA regional offices could gather information on selected indicators and make it available upon request or on some periodic basis. Alternatively, HCFA could rely on State health departments to compile and distribute the information, perhaps as a condition of Medicaid funding.

Like most other possible indicators of quality, JCAHO contingency scores have not been validated for their association with the quality of hospital care, as measured by the process or outcome of patient care. If hospitals were required to disclose the scores, HCFA and the media would have to advise consumers and the press on their appropriate use—i.e., as a guide for further questioning rather than the basis for any definitive conclusions about a hospital's quality. As discussed earlier, organized purchasers and consumer advocacy groups with experts on their staffs would be better able than individuals to interpret quality-of-care information and to exert leverage over providers to resolve any problems identified.

Option 8: Amend the Social Security Act to permit peer review organizations and the Health Care Financing Administration to disclose publicly information that identifies specific physicians.

During review of the medical care that physicians deliver to inpatients, PROs develop information related to the quality of individual physicians. Some information comes from reviewing a 3-percent random sample of the medical records related to Medicare discharges; other information comes from examining medical records selected because the PRO targets for review specific diagnoses, surgical procedures, or other areas. Although upon request PROs must disclose information that identifies hospitals, the Social Security Act forbids disclosure of information that identifies physicians (42 CFR 476.101, 104, 105, 1986 ed.). The only information identifying individual physicians that PROs and HCFA may make public is information on decisions to impose monetary sanctions or exclusions from the Medicare program.

Like the Federal Government, 6 of the 13 States that collect unique physician identification numbers (such as State license numbers) as part of their discharge data systems prohibit the release of physician-identified information.[2] To date, patient discharge data, with physicians identified, have been released only in Arizona, with no publication as yet (366). Only in Pennsylvania, where data collection began January 1, 1988, does legislation specifically mandate that data relating to the quality of individual physicians must be made available to the public (Pennsylvania Act 1986-89).

This option would change the Social Security Act to permit PROs and HCFA to make public information that identified physicians. Either in response to requests or on a regular basis, the PROs or HCFA could release, by physician, information such as mortality rates by procedure, the volume of certain services performed for Medicare beneficiaries, and the results of PRO reviews. If Congress adopted option 2, HHS could coordinate its selection of indicators to assess physicians with public release of data under this option.

Previous discussions about preserving the confidentiality of individual physicians have highlighted the conflict between public and providers' interests (451,579). Public disclosure might enable consumers and providers to identify poor-quality physicians and might prompt physicians and the hospitals where they work to improve their quality. Such disclosure, however, may also hurt the reputations and unfairly jeopardize the livelihoods of individual physicians. If physicians challenged disclosure of physician-identified data through the legal system, the judicial analysis would most likely weigh these as well as other interests.[3]

Technical problems, however, may overshadow the legal and philosophical issues. Given the current state of quality assessment, data, and the PRO process, statistics on individual physicians could mislead consumers and erroneously discredit physicians. That each physician has a much smaller number of patients and patient deaths than a hospital makes interpreting physician statistics more difficult. The deficiencies of current techniques to adjust for patient characteristics and other factors outside the providers' control apply to physician as to hospital care. But among physicians' much smaller numbers of patients, chance is much more likely than among hospitals to account for patient outcomes.

Even more basic are issues concerning the reliability of data. Current data collection systems make it difficult for researchers or quality assessors to identify all the patients treated by a particular physician. Physicians typically use different billing codes for the different locations in which they practice, and a group practice often bills for the claims of all its physicians under one code. Only 13 of 28 States with systems to collect hospital discharge data impose on hospitals unique physician identifiers (366). Attributing patients to physicians poses another thorny technical problem. The hospital designates one person as the attending physician, usually the physician who admitted the patient, but other physicians acting as consultants may play a major role in the patient's care.

---

[2] Of the seven States that do not prohibit release of physician-identified data, five have operational data collection systems: Arizona, Illinois, Nevada, Tennessee, and Washington. Pennsylvania and North Dakota have not yet implemented their systems (366).

[3] For a full discussion of the legal implications of releasing physician-identified data, see J.B. Simpson, "Release of Physician-Specific Quality of Care Information: Legal Issues" (579).

To a large extent, the validity of HHS sanctions recommended by PROs against a physician or hospital stems from the multistep process used to arrive at these decisions. This process entails review of records by trained nurse reviewers, several physician reviewers, and PRO committees; the opportunity for the physician or hospital being reviewed to attend a hearing and to provide supplementary information; and finally, review of the case by the Office of the Inspector General of HHS (see ch. 6). Although low levels of interrater and intrarater reliability threaten the validity of the PRO process in particular and peer review in general, the due process accorded to physicians and hospitals under investigation increases one's confidence in the validity of the sanctions that do result. But the findings at interim stages during the PRO/HHS process lack whatever validity is conveyed by the entire sanctioning process.

To Disseminate Information to the Public

Although much information related to the quality of medical care is already in the public domain, many individuals and organizations do not know that it is available. Available information is often not in the right format or timely enough to influence decisions. To be incorporated into consumers' choices, information must be simple and accessible when people are making decisions. Furthermore, people require skills and social support to undertake what for many is new behavior, namely interacting with physicians and raising questions about quality. The options in this section, which could be undertaken separately or together, consider more efficient ways to disseminate information on quality of care to individuals and organizations.

Option 9: Establish a new office in the Department of Health and Human Services that would be responsible for disseminating information on the quality of medical care to individuals and organizations.

As information on the quality of care becomes increasingly available, one question that arises is how to encourage the most responsible and effective provision of such information. This option would establish an office in HHS to disseminate public information on quality.

Such an office would take an active role in informing people about available information and distributing information developed by Federal programs, such as Medicare and Medicaid. These tasks would entail more than communicating knowledge. The office would inform the public about possible differences in quality among providers and interpret information from quality indicators. Over time, the office could educate consumers about the skills necessary to put their new knowledge into practice. To convey information and to engender social support for questioning quality differences, the office would work with consumer groups, medical providers, employers, third-party payers, business coalitions, States, and the media.

With such activities concentrated in one office rather than dispersed throughout HHS, HHS would be better able to apply principles of health behavior and communication in educating the public to use information on quality. Consumers' belief in the reputation of a source increases their acceptance of information. To the extent that people trust the Government as a source of health information, disseminating information through an office in HHS would increase the likelihood that people would incorporate the information into their decisions. In addition, compared with private sources, the Government would be better able to provide continuous access to quality information. Credibility and accuracy could be increased by creating an expert advisory group to review information before it is disseminated.

Congress could expand the responsibilities of the office to include all or some of those outlined in previous options, such as coordinating research on quality assessment techniques (option 1), selecting quality-of-care indicators (option 2), conducting State and local outreach activities (option 4), developing uniform data requirements (option 6), and making publicly available information on Medicare and Medicaid providers (option 7).

Combining these activities would facilitate the development and implementation of a long-term strategy regarding information on the quality of care.

Integrating and expanding these activities in a new office would require increased funding. It is also questionable whether a single office could have the experts needed to carry out such a wide range of responsibilities. Instead of creating a new office, Congress may wish to rely on existing public and private activities. Offices in HHS already undertake activities to publicize the availability of consumer information, and HCFA and the California PRO have released hospital mortality data. Several private organizations are disseminating information, and their level of effort appears to be increasing. In periodic and special publications, the Public Citizen Health Research Group has a long history of publicizing information related to the quality of physicians and hospitals. Broadcast and print media periodically gather information on indicators of quality and make it publicly available. For example, 1987 issues of *Consumer Checkbook* in Washington, DC, and the San Francisco Bay area amassed information on a range of possible quality indicators for local hospitals and analyzed how to use it appropriately. As part of their cost-containment efforts, some employers are making information related to providers' quality available to employees (253), and some private business associations are considering making information on physician and hospital quality available nationally (256). Such efforts, although limited to date, might expand as information on indicators of the quality of care becomes more generally available.

Most private groups, however, do not have the resources available to make information available for broad geographic regions, for a comprehensive set of indicators of the quality of care, or on a regular basis from year to year. As alternative approaches, the Federal Government could require State governments to disseminate such information as part of their participation in the Medicaid program or could enter into partnership with a private organization for this purpose.

A drawback of this option is the paucity of knowledge on how individuals and organizations use available information, how information can most effectively be communicated, and how existing information affects hospital and physician behavior. One would expect greater insight into these matters would permit HHS to formulate a more effective dissemination strategy than would now be possible.

Option 10: Mandate and earmark funds for research and demonstrations on methods to disseminate information on the quality of medical care.

Although one purpose of providing information on the quality of providers' care is to help consumers make more informed choices of physicians and hospitals, no empirical work has addressed whether the availability of quality-of-care information influences consumers' choices of providers. This option would fund research and demonstrations to explore the effects of quality-of-care information on consumers' decisions. These projects could be funded either instead of or in conjunction with the dissemination activities outlined in option 9.

Possible topics for research under this option include how to use the media to present information on indicators of quality so as to influence consumer choices; what type of quality-of-care information consumers find most useful in making health care decisions; what formats are most useful for providing quality-of-care information; how information learned from marketing about attracting consumers' attention can be transferred

Photo credit: American Association of Retired Persons

Little is known about how information on the quality of care can be most effectively disseminated to consumers.

to health care decisions; and other topics related to quality-of-care programs in workplace and community settings and in physician's offices. Such research would apply perspectives drawn from several disciplines and could use a variety of methods: policy review, consumer surveys, laboratory experiments, and field experiments in naturalistic environments. Initial research could conduct surveys to ascertain the level of knowledge about the quality of care among the general population and specific subgroups. Currently, this gap in information inhibits the development of effective interventions, particularly for targeted populations.

CONCLUSIONS

Although the indicators of quality examined in this report do not give conclusive evaluations of a physician's or hospital's quality, individual and organizational purchasers of care could use several of the indicators as flags, to point out areas of concern that merit further investigation. Given the current status of the indicators evaluated, formal disciplinary actions by State medical boards provide the most valid information about poor-quality physicians. In evaluating a specific physician or hospital, consumers would improve the validity of quality information if they combined the results of more than one indicator and drew information from more than a single year.

With regard to future policy in this area, those indicators that are already being used to evaluate the quality of care merit particular attention. Since governments and other entities are already disseminating hospital mortality rates, for example, the immediate task is to improve the underlying data and techniques for attributing death to prior medical care. Information on adverse events, HHS sanctions recommended by PROs, and physician specialization is also becoming generally available. Efforts to identify and improve the practices of poor-quality providers also deserve particular attention.

Although existing data sets do not allow routine evaluation of physicians' performance outside hospitals, promising efforts are underway in the United States and Canada to assess the quality of office practice across a range of medical conditions. Also promising, but not yet validated, are activities by several specialty societies to certify the competence of physicians to perform certain procedures.

Even with valid indicators of quality that are feasible to develop, using such information to guide consumers' choice of providers represents only one approach or one part of an approach to select a physician or hospital. Consumers may also rely on a primary care physician for a referral to a specialist or a hospital, a strategy that individuals often adopt. In recent years, plans that provide comprehensive care to enrollees, such as HMOs and PPOs, have institutionalized the arrangement by linking each enrollee with a primary care physician to manage that person's care. Indeed, giving consumers information on the quality of care complements consumers' reliance on a physician for referrals, because better informed consumers are more likely to be able to communicate their preferences and concerns to physicians.

Informing consumers and relying on their subsequent actions should not be viewed as the only method to encourage hospitals and physicians to maintain and improve the quality of their care. Even well-informed lay people are unlikely to have sufficient technical knowledge to judge all aspects of quality and must continue to rely on experts to ensure the quality of providers. Some experts come from within the medical community and engage in self regulation, while others operate as external reviewers through private and governmental regulatory bodies. Their continued efforts are needed for assessments of the quality of care to continue and to improve.

Chapter 2

Disseminating Information to Consumers: Present Context and Future Strategy

CONTENTS

Page
Introduction 33
The Audience for Information on the Quality of Medical Care 33
The Present Situation: Individual Consumers and Information on the Quality of Care 35
  Individual Consumers' Concerns About and Knowledge of Aspects of the Quality of Care 35
  Individual Consumers' Interest in Information on the Quality of Care 36
  Where Individual Consumers Can Find Information on the Quality of Care 37
  Reasons Individual Consumers Choose Hospitals and Physicians 38
An Effective Strategy for Disseminating Information on the Quality of Physicians and Hospitals 38
  Stimulate Consumer Awareness of the Quality of Care 40
  Provide Easily Understood Information on the Quality of Providers' Care . 41
  Present Information via Many Media Repeatedly and Over Long Periods of Time 41
  Present Messages To Attract Attention 42
  Present Information in More Than One Format 43
  Use Reputable Organizations To Interpret Quality-of-Care Information 44
  Consider Providing Price Information Along With Information on the Quality of Care 45
  Make Information Accessible 45
  Provide Consumers the Skills To Use and Physicians the Skills To Provide Information on the Quality of Care 46

Tables

Table Page
2-1. Ratings by Adults of the Importance of Selected Physician Characteristics, 1984 36
2-2. Surveys of People's Reasons for Choice of Health Services 39
2-3. Surveys of People's "Doctor-Shopping" Behavior 40

Chapter 2

Disseminating Information to Consumers: Present Context and Future Strategy

INTRODUCTION

For advice about sources of health care, Americans have traditionally relied on family or friends and on physicians. Today, most people still depend mainly on recommendations from their immediate circle of acquaintances for assistance in reaching decisions about health care providers (204,255,369,599,719) and consult with physicians for referrals to other physicians and hospitals. As changes in the medical marketplace and medical technology have increased consumers' choices and the financial importance of these choices, an issue that has come to the fore is the need for lay people to have information about the quality of care delivered by physicians or hospitals. Some observers would deny the need for such information on the grounds that the average individual lacks the ability either to make health care decisions in general or to assess the quality of physicians' and hospitals' care in particular. Consumer advocates and others who believe that better information is needed, however, do not phrase the question in terms of people's ability to judge; they simply point out that people are becoming more involved in decisions about their own health care and in making choices among providers (296).

If people are to make informed choices among providers on the basis of quality, they either must have understandable, accurate information about provider performance at hand or must be able to acquire such information easily. Until recently, information on the quality of care provided by hospitals, physicians, and other providers was not available to the public or, for that matter, to health professionals. Although quality-of-care information is increasingly being generated for public use by government agencies, consumer organizations, the popular press, and health care organizations, much of the information is unevaluated, not systematically produced and disseminated, expensive to acquire, or difficult for lay people to interpret.

The focus of this chapter is on a future strategy for effectively disseminating evaluated information to the public on the quality of physicians' and hospitals' care. As background, the discussion considers the audience for information on the quality of care and the present situation with respect to the availability of information for individual consumers.

THE AUDIENCE FOR INFORMATION ON THE QUALITY OF MEDICAL CARE

Almost all of the individuals and organizations involved in health care—employers, unions, health care providers, third-party payers, health benefit consultants, and individuals--could use accurate quality-of-care information to guide their purchase and provision of medical services. Employers increasingly are the "buyers of health care" for their employees (50), and farsighted employers are beginning to realize that quality is as important as cost in the design of benefits, purchase of care, selection of health plans, and payment arrangements between employers, unions, and health care providers (256). At least one health benefit consultant has used indicators of quality in negotiations for establishing a hospital preferred provider organization (PPO) (322).

Many unions have historically been active users of health care information when negotiating health benefits for their members. The recent trend among employers to limit employee choices to certain health care providers by limiting employees' choice of health care plans has accentuated union interest in information on quality of care. Unions, as well as employers, have little information on the quality of care provided by health maintenance organizations (HMOs), PPOs, and other types of managed care plans to which many of their members are limited (556). Validated information on the quality of medical providers in the fee-for-service sector is also scarce.

Some physicians and hospitals are ambivalent about the publication of quality-of-care information as currently constructed (41,427). Clearly, however, accurate information on the quality of hospitals and physicians could be used by physicians to select hospitals at which they will seek staff appointment; to select suitable hospitals for the admission and treatment for patients with specific medical problems, and to select hospitals or practitioners to whom to refer patients. Physicians, particularly primary care physicians, could also use information on quality to help patients choose hospitals and other practitioners. The complex nature of quality-of-care information often requires that physicians assist patients in interpreting the information's meaning.

Hospitals could use physician-specific quality-of-care information to select physicians for staff appointments and to grant admitting privileges to physicians. Hospitals could use hospital-specific and physician-specific quality-of-care information to monitor their own performance and to initiate and augment quality assurance activities and risk-management programs. Quality assurance and risk management are particularly important for hospitals in areas where providers are scarce and individuals have little choice.

Individuals and their families need quality-of-care information in order to make informed choices of physicians and hospitals. Individuals' choices are often limited. Employees' are often constrained in their choice of hospitals and physicians by the limited range of health plan options to which their employers and unions have agreed. If the only plan offered is an HMO, the employees are limited to hospitals and physicians that participate in that HMO; because of financial considerations, they would be hesitant to choose providers outside of the HMO. Medicaid recipients in some States, including California, are limited to those providers participating in Medicaid. Furthermore, millions of Americans live in areas where only one hospital or one physician trained in a certain procedure is geographically accessible. Their choice of provider is limited by geographic location. Finally, an estimated 35 to 40 million Americans are without health insurance coverage and cannot pay for care (635). These individuals are often limited in their choice of hospitals to public hospitals (72), which provide a disproportionate amount of uncompensated care (606).

Although some Americans defer decisions about choice of hospitals to their physicians, the majority of them make decisions about hospitals either alone or in conjunction with a physician. A summary of recent research found that one-third of Americans select hospitals themselves; one-third decide together with their physician; and one-third have the physician choose the hospital for them (320). Most of the decisions about which physician will provide their health care are made by individuals and their families (314). The primary health care decisionmakers within families tend to be females: women choose physicians and hospitals that family members will use as much as two-thirds of the time (320,496).

Thus, individuals' decisions are very important in the actual selection of a specific physician or hospital. Although providers and organizational purchasers of health care also have informational needs, this chapter adopts the perspective of the individual consumer in discussing both the present situation and the elements of an effective strategy for disseminating information on quality. In reading the discussion that follows, however, one should keep in mind the fact that most individual consumers' choices occur in an environment that is partly restricted by physician referral and limitations imposed by employers, third-party payers, geographic location, and lack of health insurance.

THE PRESENT SITUATION: INDIVIDUAL CONSUMERS AND INFORMATION ON THE QUALITY OF CARE

The components of a strategy for disseminating information to the public on the quality of hospitals' and physicians' care should be considered in light of several factors: individual consumers' concerns about and knowledge of aspects of quality of care, individual consumers' interest in information about quality of care, places where consumers can find information on quality of care, and reasons consumers choose hospitals and physicians.

Individual Consumers' Concerns About and Knowledge of Aspects of the Quality of Care

More than 80 percent of people in the United States have repeatedly reported that they are satisfied with the care they receive from hospitals and physicians (391,392). People's satisfaction may vary with their knowledge and rating of differences in quality. A national consumer survey found that most respondents (79.3 percent) knew that hospitals differ in their quality of care (314). Respondents with higher incomes and more education were more knowledgeable than others. Another survey reported that 69 percent of respondents deemed the quality of the health care they were receiving to be excellent or pretty good (391). People nationally expressed more dissatisfaction with the quality of care in emergency rooms and with the availability of health care on weekends and at night than with the quality of hospital care generally (390).

In rating physicians, Americans place a high value on a physician's knowledge and technical competence, but they also place a high value on the interpersonal aspects of the quality of care,[1] including the communication of information (see table 2-1). When asked the importance of certain characteristics for physicians, 96 percent of the respondents in a national survey stated that it was very important for physicians to be able to answer questions honestly and completely (see table 2-1) (392). At least three of the other characteristics rated very important by at least 92 percent of respondents pertained to clear explanations of medical problems. Having a physician spend sufficient time to diagnose and prescribe not only was rated highly, but its absence was cited as a cause of dissatisfaction by a majority of people who changed physicians. Available research on the validity of patients' assessments discussed in chapter 11 of this report suggests that people do have the ability to judge the interpersonal aspects of care.

Whether lay people have the knowledge they need to evaluate the technical competence of a provider is not entirely clear. The discussion in chapter 11 concludes that research on the validity of patients' assessments of the technical aspects of medical care is sparse and difficult to interpret. Furthermore, some research results can be questioned because experts disagree on criteria for evaluating the technical aspects of quality. In a 10-item questionnaire administered to 4,976 nonelderly persons to measure their knowledge both in choosing medical care providers (e.g., specialist v. primary care physician) and in making decisions at the time services were used (e.g., whether to have an operation), Newhouse, et al., included board certification as a valid indicator of good quality (464); as discussed in chapter 10 of this report, however, definitive evidence on the validity of board certification of the technical quality of care is lacking. Thus, depending on how one interpreted them, certain responses to the questionnaire could signify either knowledge or a difference of opinion as to the validity of the indicator as a measure of quality. Other findings of the Newhouse, et al., study suggest that consumers are knowledgeable about some matters and uninformed about others.

Bunker and Brown's study of physicians' use of medical services gives indirect evidence on lay people's knowledge of quality of care (107). Surgical rates for physicians and their wives were found to be as high or higher than surgical rates for other groups of professionals (107). The

---

[1] See ch. 3 for a discussion of the definition of the quality of medical care and its different aspects.

Table 2-1.—Ratings by Adults of the Importance of Selected Physician Characteristics, 1984

Characteristic	Very important (% of respondents)	Fairly important (% of respondents)	Not important (% of respondents)
Be knowledgeable and competent to treat your illnesses	97%	2%	—
Answer your questions honestly and completely	96	3	—
Explain your medical problems to you in a language you can understand	95	4	—
Make sure you understand what you've been told about your medical problems	95	4	—
Personally spend enough time with you to diagnose your problem and prescribe effective treatment	94	5	—
Really care about you and your health	92	7	2%
Make a special effort to get you to explain your symptoms and problems completely	92	6	1
Keep his or her medical fees reasonable	84	13	2
Tell you about steps you could take to enjoy good health such as controlling your weight, getting enough exercise, and eating the right foods	82	15	3
Have a friendly personality	70	25	4
Understand your economic circumstances	63	27	9

SOURCE: Louis Harris and Associates, *Americans and Their Doctors* (New York, NY: January 1985).

authors concluded that the physician-patient as an informed consumer places a high value on surgery and that placing a high value on surgery may overshadow knowledge about the necessity for surgical intervention. Bunker and Brown's study was done before the current emphasis on the appropriate level of care as a measure of quality. Recent findings on large variations in the use of surgical and medical procedures also have evoked interest in determining the appropriate use of services. Whether physician-patients today would act as they did in the Bunker and Brown study or whether consumers who are as knowledgeable as investigators assumed physicians to be would act in a similar fashion has not been examined.

Americans are interested in the quality of the health care they receive. Available evidence suggests that most consumers can evaluate the interpersonal aspects of health care (see ch. 11). Further research is needed, however, on patients' ability to adequately evaluate the technical aspects of care.

Individual Consumers' Interest in Information on the Quality of Care

The likelihood that an individual consumer will seek and ultimately apply quality-of-care information to choose physicians and hospitals depends in part on that person's propensity to adopt an active role in making health care decisions. National and regional surveys substantiate a willingness among some consumers, particularly younger and better educated consumers, to play an active role in making health care decisions (285). A substantial percentage of consumers actively seek and use health information in decisionmaking. A recent study of 1,833 people enrolled in Medicare Part B and State government employees enrolled in indemnity insurance plans found that just under 40 percent of respondents engaged in consumer behaviors such as seeking information, exercising independent judgment, or being sensitive to the costs of health plans (296). Younger, employed individuals were more likely than the Medicare enrollees to have greater consumer knowledge, to exercise independent judgment, and to be sensitive to cost; older Medicare beneficiaries were more likely than the State government employees to seek health information. A survey of the top 10 metropolitan areas reported that 48 percent of consumers actively acquired information and evaluated health care providers prior to using the providers' services (65).

A survey of consumers in the top 20 U.S. metropolitan areas found that 35 percent of those surveyed were very active in seeking out information and evaluating the quality of care of health care providers before using their services (65). The consumers who sought information did so because they believed that differences existed among providers. An additional 13 percent of the consumers surveyed stated that they went through the information-seeking and evaluation process when faced with an unfamiliar array of health care providers.

Anecdotal evidence suggests that few private individuals actively sought additional information about the hospital mortality data released by the Health Care Financing Administration (HCFA) in 1986. How many people knew about and then used the information in their choice of hospitals is not known, but HCFA did not receive any requests from private individuals for further information (357). Comparably, the 1986 release of hospital mortality data by California Medical Review, Inc., the California utilization and quality control peer review organization (PRO), generated only two requests by California Medicare beneficiaries to examine the primary data (435); perhaps one reason was that the costs of the information, $10 per hospital, dampened individual user interest.[2]

A sizable percentage, though a minority, of individual consumers are motivated to independently seek and use information to guide their choice of hospitals and physicians. Without strong promotional efforts to encourage other individuals to do the same, however, the effects of making quality-of-care information available may be limited. Methods of stimulating individual consumer interest in the quality of care are included as a component of the dissemination strategy outlined in the second half of this chapter.

Where Individual Consumers Can Find Information on the Quality of Care

Information on the quality of health care from sources such as the government, consumer groups, and channels including books and print and broadcast media is becoming more widely accessible than ever before to individuals and other consumers (355). Books on how to determine when to seek professional medical help and how to choose and use physicians and hospitals (64,370,563,678) have been followed by books for lay people and health professionals on how to provide and interpret useful consumer health information (150,401,512). Within the past 5 years, consumer action groups—including the Public Citizen Health Research Group, Peoples Medical Society, Center for Medical Consumers, National Women's Health Network, and the Boston Women's Health Book Collective—have offered a variety of publications with information on how to evaluate and select health care providers. Recently, newspapers and magazines have been publishing articles and publishers have been printing books that provide consumers with guidance in selecting quality medical care, both at a general level (483,542) and for specific physicians (482) and hospitals (122,277,607,693). Even some hospitals (244) and health policy organizations (498) are publishing guidelines to use in selecting physicians or hospitals.

Numerous sources now provide hospital-specific data on mortality rates[3] possibly related to the quality of care. In the early 1980s, the Public Citizen Health Research Group, a consumer advocacy organization, published a study of hospital specific mortality rates in Maryland for the 12 most common surgical procedures (55). The California PRO released mortality rate data for California Medicare patients in 1986 and 1987 (115,116), and the U.S. Department of Health and Human Services released such data for all Medicare patients in the same years (640,648). Local newspapers and magazines often report on the releases, increasing public access to the information.

In addition to hospital-specific information, some physician-specific information that relates to the quality of care is available. For example, information about formal disciplinary actions taken against individual physicians is available to consumers from State medical boards and publications (see ch. 6), and information on board certification is available from State medical societies and publications (see ch. 10).

Some health care information is specifically compiled for organizations. Health care coalitions and consortia of insurance companies provide employers, unions, and other client organizations with information on facilities, staffing, and treatment variations in various hospitals (138,416).

---

[2] See ch. 4 for a discussion of the release on information on hospital mortality rates by HCFA and California Medical Review, Inc.

[3] The individual chapters in this report that discuss hospital-specific and physician-specific indicators of quality identify sources of information on each indicator. See ch. 4 for sources of information on mortality rate data.

As part of their cost-containment efforts, employers involved in financing health care have begun to introduce consumer information programs to give employees information about the price and quality of health care. The appropriate quality of care can help contain costs for employers via decreased absenteeism, increased productivity, and decreased disability of employees (256). Burlington Industries in New York City has a program that offers employees voluntary onsite or telephone personal counseling during working hours regarding the choice of optimal health services (241). Counselors assist Burlington employees in understanding their treatment options for health problems, including what is known about the quality of various treatments and providers. As part of its cost-containment strategy, Ryder Systems, Inc., uses the MedFacts program, a computerized data base of physician and hospital profiles, to help employees choose their medical providers on the basis of quality and cost information (129).

The Washington Business Group on Health is planning a Quality Resource Center that will gather information on the quality of health care throughout the Nation (256). The center will maintain a library, a retrieval service, an 800 number, a clipping service, and online access to computerized health data bases. The center will use a variety of methods to disseminate information on the quality of care to the general public as well as to its members, including newsletters, a toll-free telephone number, articles in journals, electronic mail, reports, and seminars.

Even though sources of information on the quality of care are increasing rapidly, barriers impede many individuals' ready access to the information. Most of the information is produced sporadically and may not be at hand when needed. People may not want or be able to expend the time and money required to obtain it. Some data that are available (e.g., hospital mortality data) may be too technical for average individuals to understand. Consumers most likely to use current sources are usually people who have higher than average incomes and educational levels and are frequent users of print media (e.g., books, newspapers, and magazines) who actively seek information (617).

Reasons Individual Consumers Choose Hospitals and Physicians

Important factors in individuals' choice of hospitals and physicians are lay referrals by friends or relatives and consumers' perception of good quality care (see table 2-2). Freidson's seminal work on the lay referral system identified the recommendations of friends and relatives as central to the choice of health providers (234). Common wisdom and numerous studies support the importance of lay networks' advice on initial selection of a physician or hospital (255).

The importance of consumers' perception of the quality of care is illustrated in a number of studies (see table 2-2). Hickson, et al., found that parents' perception of a doctor's communication skills was the most important reason families had for choosing a physician to provide health care for their children (297). Accessibility and quality, as determined by recommendations of friends and physicians, were other important reasons for the choice of a physician. Stratmann found that quality of care was by far the most important of five categories (the other four are economic factors, waiting time in the doctor's office or hospital, convenience in access to care, and sociopsychological factors) in influencing the choice of health services (physician, hospital, and clinic) (603). Although Stratmann's findings must be viewed with caution because of his use of conceptually overlapping categories, a national survey confirmed his findings and reported that the key reasons for consumers' preference of a hospital were in order of importance: good medical care, proximity to home, prior experience, and a physician's recommendation (314). "Good medical care" represented a variety of responses in that survey, including availability of specialists, technology, and equipment; wide range of services offered; receiving personalized care; and good overall hospital reputation. The authors concluded that consumer perceptions of quality of care represented various components of hospital structure, perform-

Table 2-2.—Surveys of People's Reasons for Choice of Health Services

Study[a]	Population	Choice	Reasons for choice
Stratmann, 1975 (603)	521 Households in Rochester, NY	Choice of health services (hospital, physician, clinic)	Quality (> 40%), time, attitudes, cost Convenience
Flexner, 1978 (212)	Women needing abortions	Choice of an abortion service	Immediate availability of appointment Cleanliness and respectability Medical competency of staff
Glassman and Glassman, 1981 (255)	286 Women who recently gave birth	Choice of an obstetrician	Recommended by a friend or relative (46%) Recommended by a nurse (14%)
Inguanzo and Harju, 1985 (314)	Consumers nationwide	Choice of a hospital	Good medical care (48%) Close to home Availability of latest technology and equipment
Stewart, et al., 1985 (599)	229 Families in Arkansas	Choice of a primary care physician	Recommendation of friend or neighbor Personality of provider How much information provider gives Can get appointments quickly
Wotruba, et al., 1985 (719)	190 Heavy and infrequent users of care	Use of services in nonemergency situations	Heavy users: cost, convenience, physicians' interest in patient Infrequent users: lay referral, convenience, courteous staff
LeFebre, et al., 1987 (369)	241 Women who recently gave birth	Choice, of a physician for prenatal care	Professional competence (friend or physicians' recommendation, specialty, and hospital used) Convenience
Hickson, et al., 1988 (297)	750 Families	Choice of a physician for child health care	Parents' perception of their physicians' communication skills Accessibility Quality as determined by recommendation of friends or physicians

[a] Numbers in parentheses refer to numbered entries in the reference list at the end of this report.

SOURCE: Office of Technology Asssessment, 1988.

ance, and reputation, rather than any single indicator (314).

Individuals' reasons for choosing physicians for nonemergency services have been found to depend on the extent to which individuals use such services (719). Heavy users of care have been found to be most influenced by cost and third-party coverage, convenience, and the physician's interest in the patient. Infrequent users have been found to be most affected by lay referral, convenience, and courteous staff.

Studies that have examined women's reasons for selecting health care providers have found technical competence to be of importance. Important reasons for choice of an abortion service were getting an appointment right away, followed by cleanliness of the facility, respectability, and medical competence of the facility and staff (212). For women who had just given birth, two broad factors emerged as most important: professional competence or quality (as a reflection of friends' and physicians' recommendations, specialty, and hospital used); and convenience (369).

Willingness to change physicians is driven by strong motivation, except when a physician's retirement or geographical relocation is the reason. Available studies have found that the reasons that people change providers are consistent with the reasons people give when asked why they make initial choices of health providers: because of a friend's or relative's recommendation, because they are seeking better interpersonal care, or because they lack confidence in the quality of a previous provider's technical competence (see table 2-3).

Studies of consumers' reasons for choosing health services indicate that consumers often rely on the recommendations of friends and relatives in making choices of providers, in large part because of the dearth of information on the quality of care, the difficulty of evaluating the information that is available, or a belief that lay opinion is an adequate substitute for expert opinion. Available studies demonstrate that the interpersonal aspects and the technical aspects of quality are important in consumers' decisions, even when objective information about the quality of care is unavailable.

Table 2-3.—Surveys of People's "Doctor-Shopping" Behavior

Study[a]	Population	Choice	Reasons for choice
Anderson and Bartkus, 1973 (43)	579 College students in a prepaid health plan	Use of physicians outside the plan	Perceived quality of care Friends' views of quality (lay referral) Physicians' sensitivity to symptoms
Kasteler, et al., 1976 (341)	576 Families in Utah	Family member changing physician by choice without referral	Ratings of previous physicians' technical and socioemotional competence Low confidence in their physicians
Green, et al., 1979 (262)	1,278 Residents of 2 southern rural communities	Seeking new sources of primary care (not free or specialty care)	Correlates of choice: White race More frequent physician visits More shopping for acute and disabling conditions
Wolinsky and Steiber, 1982 (714)	1,530 Adults nationwide	Decision to choose a new physician	Recommendations of friends and neighbors (lay referral) Physicians' manner and personality Location, cost, and ease of getting an appointment
Marketing News, 1987 (404)	2,000 Consumers nationwide; quality minded users (largest group)	Changing health care providers	Advice of a trusted friend or relative, or recommendation of their current physician

[a] Numbers in parentheses refer to numbered entries in the reference list at the end of this report.

SOURCE: Office of Technology Assessment, 1988.

AN EFFECTIVE STRATEGY FOR DISSEMINATING INFORMATION ON THE QUALITY OF PHYSICIANS AND HOSPITALS

Information on the quality of medical care will become increasingly available over time. In the past 15 years, the volume of information available has expanded, and many signs suggest that the rate of growth will accelerate in the future. The information on the quality of care that is developed will not all be accessible to individuals; nor will it enable individuals to make wise judgments in their choice of physicians and hospitals. Some information may be untruthful or unsubstantiated; other information will be as accurate as current knowledge permits. The question is how to disseminate the latter type of information most effectively—that is, how to ensure that consumers will *acquire* state-of-the-art information and *apply* it when choosing physicians and hospitals.

The following actions are directed to achieving an effective strategy for disseminating information of the quality of physicians and hospitals. There is limited empirical evidence on how accessibility to health information affects people's choices of health care in general and whether access to information on quality of care affects people's choices of physicians and hospitals. Furthermore, a theory to explain consumer choice of physicians and hospitals on the basis of quality has yet to be developed. The strategy outlined below draws on theory and research on consumer information-processing and consumer decision-making from fields other than health and may have implications for choosing providers on the basis of quality. The specific components of the strategy are unproven and would require empirical verification before adoption.

Stimulate Consumer Awareness of the Quality of Care

Before making choices, consumers must perceive differences in the product or service and the possibility of making a choice (198). Most consumers recognize that there are differences in quality among providers (315), and a sizable minority are motivated to seek and use information on quality to guide their choice of physicians and hospitals (65). Consumers in the latter category are predominantly white, have high incomes, and are well-educated (243,315,341).

To enlarge the audience for quality-of-care information, an initial step would be to make consumers aware that there are differences in providers' care. Individuals who cannot envision the possibility of an option do not consider alternatives but exercise their habitual preferences (145,198). Informing consumers who do not already know it that hospitals and physicians vary in the quality of care they provide could stimulate greater efforts by consumers to acquire and use quality-of-care information in choosing providers.

In addition to a lack of information, psychological factors, which are difficult to overcome, may blind individuals to possible options or allow them to see alternatives only if they are presented in certain ways (619). Some potential choices may never get considered because an individual's habitual ways of framing preferences may exclude them. Since there are few data in this area, more research is needed before framing theory can be applied to choosing providers on quality grounds.

For some consumers, improved knowledge about differences in the quality of care among providers and the accompanying perception of the risk posed by poor care may increase their interest in quality-of-care information. The greater the potential harmful or undesirable effects of using a product, the higher the perceived risk and the greater propensity to seek out more data (60,198). Perceived risk can be equated with a sense of personal susceptibility (63), for example, the belief that one may be at risk when receiving medical care. Most people do not feel themselves at risk when receiving health care services in general (391). Medical care is not a homogeneous commodity, however, and individuals seeking treatment for serious conditions may have a greater sense of personal susceptibility than individuals seeking care for minor ailments.

Provide Easily Understood Information on the Quality of Providers' Care

Numerous factors affect people's ability to understand information. In general, there are limits on people's ability to process information (431,577). Even for individuals whose information-processing abilities are high, information needs to be easy to understand, because processing information requires the expenditure of finite resources (primarily effort and time) (70) that individuals may not want to expend. New information is especially difficult to process, because a person attaches meaning to a message by comparing it with old information stored in memory (198). For most people, quality-of-care information will be new, particularly if specific indicators of quality rather than general statements about quality are presented. Consequently, care must be taken to disseminate meaningful quality-of-care information that is easily understood.

Furthermore, language will pose a barrier for some consumers. About 11 percent of the U.S. population speak a language other than English at home (634). To reach these individuals, information on the quality of providers' care will have to be translated into languages other than English; alternately or additionally, cultural interpreters may be needed.

To more effectively inform consumers about the quality of providers' care, limiting information to only a few indicators of quality will probably be necessary. People can consider only a few items at any one time (431,577). Information is processed as a unit or chunk—a person's processing capacity has been estimated as being anywhere from four to seven chunks (198). Research on label formats that describe the nutritional content and quality of food products suggests that when information is given about numerous attributes, consumers find the labels difficult to understand (633). Most food choices, however, are made at the time of purchase, whereas, except in emergencies, most health care provider choices are made before an encounter.

Factors specific to an understanding of technical topics will also affect a strategy for informing consumers about the quality of medical care. People vary considerably in their understanding of information about medical details (202). Understanding is diminished by the use of medical terminology and by the use of common English terms that have special medical meanings (e.g., history, acute). Some individuals have no or little knowledge against which to interpret the information presented (565).

Some consumers may find information on the quality of care as difficult to understand as medical terminology. Terms such as mortality rates and iatrogenic illnesses are technical words that are not employed in everyday life. Other terms used to designate quality indicators, such as volume of services and scope of services, are common words but they have a special significance as potential indicators of quality. To a lay person, the phrase scope of hospital services suggests the specific services a hospital offers its patients. As a quality indicator, scope of hospital services refers to a hospital's resources for the medical conditions it professes to treat, or resources for the medical condition affecting a potential patient (see ch. 9).

Information would be more intelligible to more consumers if the use of technical terminology and the use of terms with special medical meaning were limited and words used in everyday language were substituted. The term hospital-acquired infection might be used instead of nosocomial infection. Words used frequently in everyday language are more easily comprehended and remembered than words used rarely or not at all in everyday conversation. The most suitable language of the information will probably vary by consumer groups because of differences in culture and educational level.

A particular problem is communicating information to consumers about mathematical concepts such as risks, percentages (202), and probability. Understanding the data on some quality indicators, including hospital mortality rates, requires an understanding of probabilities and risks. Because of the problems many people have in processing mathematical concepts, errors and exaggeration of risks occur in making choices (619). One way to increase comprehension might be to use both numeric and nonnumeric terms (such as small and large) to describe probabilities and risks; also the meaning of small and large in other and more familiar circumstances could be described.

Finally, the manner in which risk information is formulated can influence people's choices (337). Empirical studies of how the formulation of information affects choosing between medical interventions show that the choices differ by whether probabilities are formulated in terms of survival or of death.

Present Information via Many Media Repeatedly and Over Long Periods of Time

Sources of information vary among individuals and situations. Furthermore, people making choices use a variety of sources, usually in combination, in their search for information when making choices (145,198,541). Although lay referral may remain as one of the most important sources of information for individuals when choosing health care providers, they nevertheless would benefit from access to a number of alternative sources. As an example, the most effective self care programs—the choice being self-care or physicians' care—have used more than one approach to provide information, including written material, group education sessions, and individual counseling (253). Special outreach efforts and information tailored to various educational levels have been necessary to ensure that these programs reached lower socioeconomic and minority groups.

There are a variety of media that can be used to convey information, and one form may be better than another for conveying certain aspects of information (198). The mass media (print and electronic) inform average consumers about matters, such as the availability of products and services and the features of particular brands (145,541). The print media are probably consulted more than the electronic media for choices that involve a high degree of personal concern and have serious consequences (281). In addition, the effectiveness of a particular medium depends upon the type of consumer. In general, better educated consumers tend to rely more on the print media than do other consumers (198). A recent survey reported that printed materials, television, and informal networks of lay people and professionals were the most frequently used sources of information for respondents. Few respondents reported receiving health information from radio organizations (145).

Messages need to be repeated over a long period of time because people have limited ability to retain information (198), either because the memory of the message fades with time or other information interferes with retrieving the information (200). People's retention of quality-of-care information specifically appears to be slight (367). A survey of clients found that 2 months after the widely publicized release of hospital mortality data by HCFA, 48 percent of 900 interviewees in Milwaukee, Wisconsin, recalled that they had read articles or heard news reports on the topic, but only 6 percent accurately recalled the content of the message. Also, the probability that the message will be processed and used in making a choice is determined in part by attitude and by social and situational factors (210). If information on the quality of care is presented only once or twice, a person may not be interested in it at the time it is presented. A sudden loss of employment and loss of health insurance coverage, for example, may cause an individual to ignore the information if he or she intends to delay a scheduled elective surgery.

Present Messages To Attract Attention

Capturing an individual's attention may not necessarily lead to the person to acquire and use the information presented, but it is a step in that direction. Capturing attention is influenced by individual characteristics. As noted earlier, one reason for repeated presentations of the same message is that people pay attention to messages that are relevant to their needs. People also try to maintain a consistent set of beliefs and attitudes (422) and attend to messages that enhance consistency and avoid information that challenges it. Thus, some individuals have to be sensitized to the fact that medical providers vary in quality of care and that they can choose among providers.

Another major factor in capturing attention is the characteristics of the message. How attributes such as size, color, intensity, contrast, position, structure, and movement affect the ability of information to attract attention has been well researched in the marketing field (198). Although consumers' choices of hospitals or physicians are rarely on-the-spot decisions, the lessons from marketing could be applied to disseminating information about the quality of such providers' care.

Present Information in More Than One Format

People use complex information processing strategies to choose among alternatives that differ on many features. One approach to processing information is to evaluate all the features of each alternative; another approach is to evaluate all the alternatives with respect to a single feature, then a second feature, etc. (70). People require less effort to process information in the former way than in the latter.

Information on the quality of care provided by physicians and hospitals could be presented either by individual physician and hospital or by characteristic across physicians and hospitals. In the former case, the characteristics of individual physicians and hospitals could be displayed with respect to quality indicators (e.g., the specialty status of the physician, the presence or absence of disciplinary actions, the mortality rates of a hospital, and the scope of services of a hospital). In the latter case, quality indicators could be arrayed with the comparative standing of individual physicians and hospitals listed under each indicator.

Presenting information on the quality of providers' care calls for both approaches, because consumers have different levels of knowledge. Consumers who are thinking about going to or continuing to go to a particular physician or hospital would probably prefer to choose by the characteristics of the particular physician and hospital they are considering. Other consumers might prefer information presented in a format designed for comparative choice among several physicians or hospitals. Similarly, consumers with limited time would prefer to have information about a particular physician or hospital, while those with more time might accept information arrayed for comparative choice.

Use Reputable Organizations To Interpret Quality-of-Care Information

Consumers believe that reputation is a good proxy for quality, particularly when they find it difficult to judge quality and therefore perceive their choices as involving a high level of risk (60). Reputation of the manufacturer is often used as a proxy for quality in the choice of over-the-counter drugs, such as aspirin. Many consumers choose providers on the basis of their belief that reputation indicates quality. Providers' services involve some intangible characteristics (373,671), and the difficulties inherent in evaluating such characteristics may be a problem for consumers. This problem may lead consumers to rely heavily either on a provider's reputation as known either directly or through recommendations from friends (609). Indeed, lay and professional referral, the most common sources of information on the quality of providers' care, are based mainly on providers' reputations. Consumers' acceptance of physicians' selections of hospitals (320,496) and referrals to other physicians illustrates consumers' belief that physicians are qualified to evaluate medical care.

A specific aspect of reputation is the credibility of the source of the information and consumers' trust and belief in the source's ability to evaluate the reputation of the provider. To ensure accuracy of information and to obtain public confidence, the source that interprets the information on the quality of care provided by physicians and hospitals should be a reputable one. Consumers' belief in a source of information increases their acceptance of the information. Trusting the source simplifies their decision; they can discontinue their search for information if the information they need has been acquired and processed by trusted regulators or consumer groups. The same source could then disseminate the information on providers' quality to other media and directly to consumers.

Consider Providing Price Information Along With Information on the Quality of Care

At times people's beliefs are inferential (210). Some people, for instance, believe that if the price is high, the quality is good (198). People tend to rely most heavily on price cues when quality information is unavailable and when they have little experience in evaluating the product (or service) (437,575). Indeed, in assessing health care providers, particularly hospitals, patients often use price as a surrogate for quality (407).

In some cases, consumers go beyond quality when choosing providers; they make price/quality trade-offs. When making such trade-offs, consumers require price information. Consumers have a fairly great amount of information about the prices for routine care, but less about prices for surgical care (407). The reason may be that obtaining information about frequently used medical services cost less than obtaining information on other types of medical services (407). Another possibility is that consumers may be more interested in the price of services that are usually not covered by insurance (e.g., pediatric care and routine checkups) than in price information for services extensively covered by insurance (e.g., surgical services).

Make Information Accessible

Consumers seek to process as little information as possible in order to make rational decisions quickly (268), and once they find a satisfactory alternative, they will discontinue their search rather than searching until they find the best alternative. The ease of obtaining information is an aspect of accessibility that is important to consumers when making decisions about providers of health care. Consumers are more likely to obtain and use information if it is accessible at all times and if the physical location of the source of information is where the consumer can reach and use the information with the least possible expenditure of time and energy. Financial access to information is also important to consumers. The costs of information and the way information is provided should not deter consumers from seeking it. Making accurate information easily accessible improves the chances that consumers will use accurate information rather than poor information in making their choices.

Access to information on the quality of providers' care has been growing concurrently with the availability of such information. It appears that employers and the public increasingly will have information about indicators of quality of care accessible to them (442).

If information is to be effective, it must be accessible when consumers make decisions about providers, when they are changing providers, and

Photo credit: American Association of Retired Persons

Consumers are more likely to obtain information if the physical location of the source is easily accessible, such as in senior citizen centers.

when they are considering a physicians' referral to a physician or hospital. People search for information from sources that are easily accessible, in location, time, and monetary costs, and they continue their searches longer when the sources of information are accessible than when the information is hard to obtain.

Releasing new information through multiple forms of the mass media increases its accessibility. The release of hospital mortality statistics by HCFA is a step in that direction. Those statistics were reported not only in the print media, but also on the radio and television (see ch. 4). Another step might be to make quality-related information continuously accessible to consumers in hard copy and through computer terminals in libraries, senior citizen centers, adult education centers, community centers, and other facilities. Hard copy information could be provided to physicians, particularly referral physicians; this would assist them to make wise referral choices and to help patients who want the information interpreted. Cable television exposures could be considered as could "hot lines" that could provide a source of continuous information.

The acceptance of information on the quality of a providers' care is increased when it is accessible in familiar settings, such as libraries and senior citizen centers, where needed social support is present. Studies of consumers' reasons for choosing health services indicate that consumers often rely on the recommendations of friends and relatives; lay opinions and social networks play an important role in the evaluation and decision processes regarding choice of physicians and hospitals. Consumers need social support from peers, family, and friends in making choices of health providers. Expert-based information may seem less foreign if it is presented in familiar settings.

Social support helps reinforce a behavior change. The sources of reinforcement, which include family, peer groups, teachers, employers, health providers and the media, vary with the change being considered (262). The particular groups needed for some choices have been identified. A review of 150 articles on nutrition found that people need not only information but also support and followup reinforcement from family, friends, and primary care physicians in making choices about nutritional intake (252). Furthermore, the relative importance of particular support groups has been established for a few behaviors in certain settings. Adolescent drug-taking behavior, for example, is most influenced by approval from friends (321), especially a best friend (338a). Sources of support when making choices about providers on quality grounds and their relative importance are other areas that need to be examined.

Provide Consumers the Skills To Use and Physicians the Skills To Provide Information on the Quality of Care

Specific skills are required for consumers to be able to *use* effectively information on the quality of care that they have acquired. Knowledge alone is not sufficient. If the purpose of providing information is to change health behavior, certain knowledge about how to follow the physician's advice is essential (62). If the purpose of providing information on indicators of quality is to assist consumers in choosing physicians and hospitals, consumers will need skills or assistance in interpreting the information and in asking questions about its significance in individual situations. Physicians are likely sources of such information.

Consumers who call on their physicians for assistance in interpreting the meaning and use of indicators of the quality of care need skills to ques-

Photo credit: American Association of Retired Persons

Consumers need the skills to enable them to ask their physicians the right questions about their conditions and treatments.

tion them. Although some consumers are hesitant to question physicians, two experimental studies demonstrate that patients can successfully be "coached" to ask more questions of physicians and to secure more information about their conditions and treatments (264,540).[4] Consumers need the skills to make them capable of asking the right question. In addition, physicians must be willing and able to provide help and interpretation. Some physicians might benefit from continuing education to make them aware of their patients' desire for information and to acquire the skills and resources to answer their patients' questions. Physicians need skills to ensure that the desired information has been transmitted.

---

[4] Some organizations have started to provide information to consumers on how to ask questions of physicians, e.g., the National Women's Health Network has a publication "Plaintext Doctor-Patient Checklist," which lists questions to ask physicians during an appointment (458). A publisher, Krames Communications, issues a comic-book format brochure, "Asking Questions: For Only the Best Health Care," with types of questions for patients to ask physicians during different types of encounters (358).

Chapter 3

Evaluating Quality From the Perspective of Individual Consumers

CONTENTS

Page

Introduction 51

Defining the Quality of Medical Care 51

Framework for Assessing Quality of Medical Care 55

  Progression of a Person Through the Spectrum of Medical Care 55

  Approaches to Assessing Quality 57

  Aspects of Medical Care To Evaluate 58

  Possible Indicators of the Quality for Individual Consumers 59

Indicators of Quality Selected for OTA Evaluation 61

  Criteria for Selection 61

  Indicators Selected for Evaluation 66

  Evaluation of the Indicators: General Issues 66

Figure

Figure *Page*

3-1. Progression of a Person Through the Spectrum of Medical Care 56

Tables

Table *Page*

3-1. Aspects of Medical Care To Evaluate 58

3-2. Possible Indicators of Hospital Quality and Their Relationship to Aspects of Medical Care 60

3-3. Possible Indicators of Physician Quality and Their Relationship to Aspects of Medical Care 61

3-4. Distribution of Office Visits to Physicians, by Physician Specialty and Patient Age, 1985 62

3-5. Management of Specific Conditions as Possible Indicators of Quality 63

3-6. Considerations in Selecting Indicators of Quality for OTA Evaluation 64

3-7. Issues Addressed by the Indicators Selected for OTA Evaluation 65

Chapter 3

Evaluating Quality From the Perspective of Individual Consumers

INTRODUCTION

For some time, physicians and other medical professionals have assessed the performance of their peers. From Florence Nightingale in the field hospitals of the Crimean War to E.A. Codman in surgical wards of Boston during the early twentieth century and Osler Peterson among general practitioners in North Carolina after World War II, medical professionals motivated by a deep concern for their patients' welfare have strived to measure the quality of medical care so that providers could improve it. Along with medical professionals, concerned people from fields such as statistics, politics, and religion have pioneered techniques to evaluate the efficacy and safety of technologies, and, in turn, the quality of medical care (628).

Quality assessments have customarily taken the perspective of the medical provider. Recent events, however, have promoted consumers' role in evaluating providers and making decisions about medical care. Efforts to advance consumers' interests are occurring throughout society, and the changing role of consumers within medical care reflects this societal trend. The increased emphasis on consumers also reflects the influence of strategies to increase price competition in medical care. People have always had a legitimate interest in the quality of their medical care. But recent policy changes have created a milieu in which the consumers and providers of medical care have become more sensitive to price. In that milieu, information about the quality and cost of care is needed by consumers to aid them in selecting physicians and hospitals.

Given that context, it is important to examine the perspective of individual consumers on the quality of medical care. Do consumers' needs and concerns differ from those of medical providers in ways that should be taken into account in the design and content of quality assessments? This chapter explores that question. The chapter first develops a definition of the quality of medical care that incorporates its many dimensions. In a section presenting a framework for assessing quality from an individual consumer's perspective, the chapter describes the progression of a patient through the spectrum of medical care. Then it discusses approaches to assessing quality and aspects of medical care that affect health and patient satisfaction and presents possible indicators of quality. The chapter concludes with a discussion of the indicators selected for evaluation in this report.

DEFINING THE QUALITY OF MEDICAL CARE

Like other intangible concepts, the quality of medical care is difficult to define. Indeed, quality acquires concrete properties only when one measures it. But attempts to define quality in the medical field are plagued not only by the abstract nature of quality but also by particular characteristics of medical care.

Medical care is intended to promote, maintain, and restore health (186). Although the purpose of medical care is to help patients, appropriate care and desirable outcomes vary tremendously depending on the individual patients' circumstances. Healthy infants require immunizations to prevent once-common childhood diseases and ultimately to lengthen their lives. Screening during infancy and adulthood may detect conditions that treatment can correct or ameliorate. Throughout life, treatment may cure acute conditions and relieve the symptoms of chronic ones. Medical care may also help people deal with their physical and emotional problems. For people facing death or intractable conditions, medical care may offer palliative measures that reduce suffering and help people to die with dignity. Thus, the appropriate content of medical care stretches from the prevention of illness to diagnosis, rehabilitation, counseling, and other therapy, and desirable outcomes of care range from reduced illness, deterioration, and pain to increased longevity, mobility, and emotional well-being. And all of the activities and outcomes of care presume that people seeking care, especially in emergencies, promptly reach providers who can manage their conditions.

To a large extent, the diversity of acceptable outcomes for patients reflects the many dimensions of health. According to the definition of health adopted by the World Health Organization: "Health is a state of complete physical, mental, and social well-being and not merely the absence of disease or infirmity" (718). This definition stresses the positive aspects of health while incorporating the notion that health relates to physical functioning, mental health, and social functioning.

Noting the complexity of medical care, prominent scholars have stressed the importance of evaluating both its technical and interpersonal aspects. Technical care is the application of medical science and technology to a problem; and interpersonal care or the art of care refers to the personal interaction between patient and medical care provider (105,183). In practice, the technical and interpersonal aspects of care are intertwined; sensitivity and caring enter into technical care, and technical expertise is part of interpersonal care. Both these aspects deserve attention in evaluations of the medical care that patients receive.

Besides taking into account the many dimensions of medical care and health outcomes, a definition of the quality of medical care must recognize the limits and continuing evolution of medical knowledge. Medical knowledge and its application in medical technology cannot guarantee improvement in a patient's health. At best, medical care applied appropriately can improve the likelihood that a patient will get better. Rarely is a medical technology 100-percent efficacious. The use of medical technology also carries some risk, and this must be weighed against the likely benefit.

Photo credit: March of Dimes Birth Defects Foundation

Technical and interpersonal aspects are intertwined in medical care, such as the rehabilitative therapy shown here, and both deserve attention in evaluations of the quality of care.

The probabilistic nature of patient outcomes flows from the variation in patients, providers, and environments. Even medical technologies found to be efficacious for treatment populations in the ideal circumstances of randomized clinical trials may not benefit a particular patient. Patients' physical and emotional conditions differ in ways that affect treatment results, and these differences may be unknown or unpredictable when medical decisions are made. Another point relevant to the quality of hospitals and physicians is that providers themselves vary in ways that may affect what happens to patients' health.

In a larger sense, the uncertainty surrounding patient outcomes stems from the fact that medical care is but one influence on the health of an individual or a population. In fact, an individual's genetic makeup, environment, and lifestyle seem to play a greater role than medical care in explaining the causes of death and illness that now predominate in the United States.

What is considered appropriate care evolves with advances in medical science and technology. As knowledge continues to expand, some technologies (e.g., gastric freezing for ulcers) become obsolete and should be discarded, and others (e.g., cimetidine) are shown to be efficacious and should be adopted.

Over the years, scholars have taken many different approaches to incorporating these complexities into a definition of the quality of medical care. Quebec's Commission of Inquiry on Health and Social Services (the Castonguay Commission) refused to define quality and commented that "...choosing among the possible definitions of the quality of care leads to rejecting part of reality and to reducing the meaning of quality to one or some of its dimensions" (505). Rather than defining quality, the commission identified how perspectives on quality differ: Producers evaluate technical aspects of services, mostly for care of the sick, but pay scant attention to access or distribution of care; consumers wish a minimum level of technical competence but emphasize more heavily ease of access, continuity and humanization of care, and prevention of disease; and society, from another level, focuses on how care affects the population's health and how the social and economic efficiency of the system conforms to society's priorities.

In a similar vein, Donabedian acknowledged the different views of providers, consumers, and the overall society: Physicians have usually confined their evaluations to technical performance, patients have shown more sensitivity to how they are treated, and society has had more interest than individual providers or consumers in the equitable distribution of medical care and the public health benefits of care, such as prevention of communicable disease (186). But Donabedian also stressed that all view both the technical and interpersonal as important (183).

Donabedian's discussion culminates in ". . . a unifying concept of the quality of care as that kind of care which is expected to maximize an inclusive measure of patient welfare, after one has taken account of the balance of expected gains and losses that attend the process of care in all its parts" (183). To the extent that the patient bears the cost of care, Donabedian includes cost in this concept of quality on the grounds that one may add cost, as an unwanted consequence of care, to expected risk in assessing the patient's net benefit. However, Donabedian keeps accessibility, the ease with which care is initiated and maintained, separate from quality.

Although it was not developed specifically for quality assessment, Palmer has used an Institute of Medicine definition of a quality assurance system that also refers to resource constraints: "The primary goal of a quality assurance system should be to make health care more effective in bettering the health status and satisfaction of a population, within the resources that society and individuals have chosen to spend for that care" (475).

Another definition stresses the response to needs and defines quality as "the degree to which health care needs (educational, preventive, restorative, and maintenance) of an individual or group are identified in an accurate, complete, timely manner, and the resources (human and other) necessary to meet these needs are applied in a timely manner and as effectively as current knowledge allows" (524).

This OTA report examines several possible indicators of the quality of care provided by hospitals and physicians, not the quality of care of a managed health care system or the quality of the entire U.S. health system. Reflecting this task and the points discussed above, the report uses the following definition of quality to guide the discussion:

> The quality of a provider's medical care is the degree to which the process of care increases the probability of desired patient outcomes and reduces the probability of undesired outcomes, given the state of medical knowledge.

Under this definition, medical care consists of the technical and interpersonal interventions that providers apply to improve patients' health and satisfaction. The quality of medical care delivered by a hospital or physician is judged by the likelihood that the care will achieve the patient outcomes desired, and this likelihood depends on the relationship between certain medical practices and the effects on patients. Desired and undesired outcomes, comprising positive and negative effects, relate to the many dimensions of health and patient satisfaction. Which ones predominate varies with the individual patient or condition.

The definition of quality of care used in this report incorporates some, but not all, aspects of people's access to care. A host of factors—psychological, physical, social, and economic—determines whether a particular person decides to seek care for a medical condition. All of these factors relate to the accessibility of care to an individual (i.e., the ease with which a person can gain entry into the medical care system). One important factor is the cost that the person expects to pay, which in turn depends on insurance coverage (or the lack of it) and the provider's charges (386,463).

Although the choice of health insurance coverage and the decision to seek care wield great importance, scholars have usually separated issues of access from those of quality, and this report generally follows that convention. But two aspects of access overlap with quality and have such strong implications for patient outcomes that they are included in this report: providers' responsiveness to urgent or emergency care and providers' referral of patients to the appropriate level of care. Even after a person decides to seek care from a specific provider, barriers may prevent the person from obtaining care or from reaching the appropriate level of care. At the same time, the responsiveness of hospitals and physicians, especially to urgent or emergency situations, may well affect the person's eventual health outcome. The procedures of a hospital or physician may keep the patient from seeing a health professional in a timely manner. A hospital emergency room that transfers a patient in an unstable condition to another institution because the patient lacks insurance may jeopardize the person's health. On the other hand, failure to transfer a high-risk mother or baby to an institution with a higher level neonatal intensive care unit may also jeopardize health.

Most hospitals and physicians practice independently and typically do not assume responsibility for a clearly defined population. It would not be reasonable to hold these providers responsible for the ease of access perceived by all the people in a certain area, even if barriers had impeded people's access to care and harmed their health. Physicians and hospitals operating as separate units have not had the same responsibility for ensuring that certain facilities and personnel are available as health care systems, such as prepaid group practices. On the other hand, hospitals and physicians have a core group of people who rely on them for care. Once that relationship has been established, it seems reasonable to hold providers responsible for making their services easily accessible to these patients. Moreover, it would be reasonable to include issues of access in evaluating the quality of a health care plan that assumed responsibility for a given population and the quality of a national health care system, which bore responsibility for the country's population.

Excluded from this report's definition of the quality of care are considerations of cost and efficiency. Conceptually, medical care's effects on patients' health and satisfaction differ from its effects on costs. Even more important, however, when making decisions about medical care, consumers, providers, and policymakers weigh the likely health benefits against their costs. Costs indicate what people must forgo in other goods and services in order to obtain the health outcomes that they desire. Indeed, behind recent changes in payment policies has lain the intention of heightening the cost consciousness of consumers and providers who make decisions about using medical services. From a policy perspective, separating cost from quality or health effects permits analysts to monitor any changes in health that occur as costs change and to identify what is being gained or lost. Such information also permits one to evaluate the efficiency of the provider, in this case the use of resources (costs) to achieve a given level of health benefits.[1]

---

[1] In spite of the conceptual distinction between cost and health effects or quality, it is unlikely that peer reviewers will incorporate the distinction into actual assessments of providers' performance. Either implicitly or explicitly, quality assessors develop indications for the appropriate use of a certain procedure, such as coronary artery bypass surgery, or identify medical interventions deemed necessary to manage a particular diagnosis. With the increased cost consciousness in the U.S. medical community, peer reviewers most likely will factor cost as well as health effects into their criteria.

Also excluded from the definition of quality in this report are amenities that may be provided in the course of medical care. What sets the activities that are considered medical care apart from these other areas is that medical care is undertaken expressly for the purpose of affecting health. Although amenities such as office furnishings and hospital food certainly influence patients' satisfaction, in keeping with this interpretation of medical care, this report excludes such amenities because their main purpose is not to improve health status (201).

In addition to people who receive medical services, many individuals and organizations are consumers of medical care in the sense that they make decisions about purchasing such care. Parents arrange for the care of their children, and grown children may arrange for the care of their elderly parents. Third-party payers, both governmental and private, decide which services are covered, under what circumstances coverage applies, and how much will be paid; insurers may also contract with selected providers. In the workplace, employers and unions make many such decisions that affect the availability of workers' medical care. In addition, public interest groups and associations of particular types of consumers, such as elderly people, represent the interests of individuals in policy decisions. And all of these organizations provide information that is intended to help individuals choose medical providers.

In constructing a framework to assess a medical provider's quality, this report takes the perspective of the individual consumer. This restriction reflects the fact that medical professionals provide care to benefit individuals. As discussed in chapters 1 and 2, however, the perspectives of both individual and organizational consumers are clearly germane to the feasibility of using certain indicators and to the policy implications of publicizing information on quality. The report therefore considers both organizational and individual consumers in its sections on feasibility and policy implications.

FRAMEWORK FOR ASSESSING THE QUALITY OF CARE

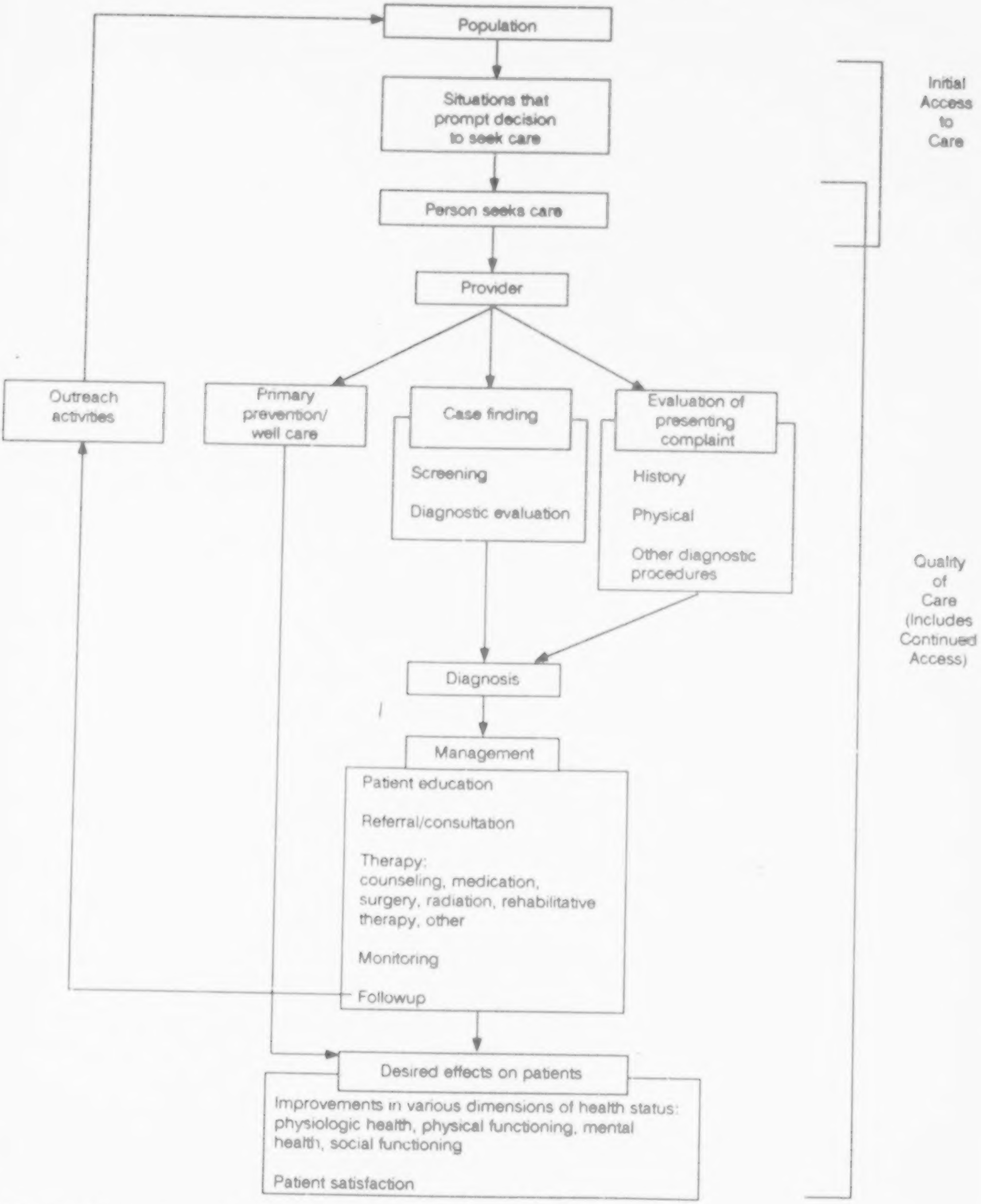
Progression of a Person Through the Spectrum of Medical Care

A framework for individual consumers to assess quality should address the choices that people face and the care that they receive as they enter and proceed through the medical care spectrum during an episode of care. Figure 3-1 describes the key elements in the progression of a person through that spectrum.

The population consists of the people who may use a particular provider for medical care. For a hospital or physician within a prepaid group practice, the enrollees of the group comprise the population at risk. Enrollees are covered for care in the group's facilities and, presumably, will use the group's providers in most circumstances. By comparison, most hospitals and physicians in the United States have a population that is much less well defined. A given hospital may draw most of its patients from a certain area, but people from other areas or their physicians may also prefer that hospital and use it for hospital care. The same situation applies to physicians who provide care on a fee-for-service basis. Especially indistinct is the population of a specialist or subspecialist (e.g., a radiologist or neurosurgeon) who obtains patients primarily though the referrals of other physicians. Even physicians in an individual practice association (IPA), a type of health maintenance organization (HMO) in which physicians continue to practice separately but agree to provide covered services for a monthly per capita payment, do not have a defined population for whom they are responsible. IPA enrollees, like others who pay fees for services, may choose their physicians from several who participate in the plan.

As shown in figure 3-1, conditions arise that prompt people to seek medical care. As noted earlier, many factors influence the decision to seek care and the ease with which people obtain appropriate care. Of key importance for evaluating the quality of medical care are how providers respond

Figure 3-1.—Progression of a Person Through the Spectrum of Medical Care


Population
Situations that prompt decision to seek care
Person seeks care
Provider
Outreach activities
Primary prevention/ well care
Case finding
Screening
Diagnostic evaluation
Evaluation of presenting complaint
History
Physical
Other diagnostic procedures
Diagnosis
Management
Patient education
Referral/consultation
Therapy: counseling, medication, surgery, radiation, rehabilitative therapy, other
Monitoring
Followup
Desired effects on patients
Improvements in various dimensions of health status: physiologic health, physical functioning, mental health, social functioning
Patient satisfaction
Initial Access to Care
Quality of Care (Includes Continued Access)

SOURCE: Office of Technology Assessment, 1988

to people attempting to obtain care, especially in urgent or emergency situations, and whether people reach the appropriate level of care. Issues of access with quality implications arise not only when a person initially seeks care during an episode of illness, but also when a person tries to return for followup care or to pursue referral services.

The middle part of figure 3-1 illustrates the different components of medical care. If a person seeks care for a specific complaint, the physician should obtain relevant information from the patient, perform an examination, and conduct any appropriate tests needed to make a diagnosis. Whether a person seeks care for a particular problem or for a checkup, the physician should follow certain procedures to screen for the presence of certain chronic conditions (e.g., taking the patient's blood pressure to detect hypertension) and to prevent the occurrence of disease (bringing immunizations up to date).

In many of these steps, the physician or other health professional requires more than physiologic and physical information. To evaluate and diagnose a patient's condition, the provider must often know the patient's psychological state; lifestyle; and environment, including working conditions and social interaction with family and friends. Whether the provider can elicit such information depends on the relationship that the provider has established with the patient.

The pervasiveness of the patient-provider relationship and its importance for many aspects of medical care are evident as one proceeds beyond diagnosis to the management of a patient's condition. Developing a strategy to manage the patient's condition requires that the physician know the patient's preferences and goals. For example, appropriate therapy for an orthopedic injury in a professional athlete may well differ from what would be appropriate for someone less interested in athletic competition. Whether to seek a consultation from another physician or to refer the patient for more specialized care may also depend on the patient's preferences and goals. The relationship established with a patient would be expected to have major importance in any situation in which a physician was trying to persuade a patient to engage in certain behavior—in counseling the patient about prevention, a chronic condition, medication or other regimens, rehabilitative therapy, and followup care.

As figure 3-1 indicates, medical care is intended to maintain or improve patients' health status across a wide range of dimensions and to satisfy patients. In some cases, medical care can improve a condition by curing disease, alleviating symptoms, arresting disease progression, restoring function, or reassuring a person who is worried but well. Medical care may also benefit a person whose condition cannot be improved if the provider can clarify a situation and reduce uncertainty. Because of the many factors besides medical care that influence health and satisfaction, even the most effective medical care provided in the most sensitive way may not result in the outcomes desired. Nevertheless, situations of different kinds prompt people to seek medical care, and patient satisfaction and health improvements are the intended results.

Approaches to Assessing Quality

The quality of medical care can be assessed by evaluating the structure, process, or outcome of care (183). Each of the approaches in this commonly used schema focuses on the measurement of quality at different points in the spectrum of medical care.

The *structure* of medical care subsumes the resources and organizational arrangements that are in place to deliver care. Structural characteristics used in assessing quality include the number, type, and distribution of medical personnel, equipment, and facilities. The presence of a quality review committee; procedures for coordinating nursing and other services; and organizational arrangements of physicians, such as solo or group practice, also relate to structure. Behind using structural characteristics to assess the quality of care lies the assumption that such characteristics increase or decrease the likelihood that providers will perform well. This assumption in turn raises the issue of whether specific structural characteristics of medical care are in fact associated with better performance or process.

The *process* of care refers to the activities of physicians and other health professionals in caring for patients. Assessing that process entails evaluating the performance of the different aspects of care considered important. The content of appropriate care evolves over time as science and technology progress and as consumers change their expectations of technical and interpersonal aspects of care. Although procedures to be followed may be specified by medical condition, what is appropriate under each aspect ultimately depends on the particular patient.

The major difficulty with assessments of process is the dearth of information about the efficacy of most medical procedures. It is reasonable to judge providers' performance only in relation to procedures likely to improve or harm patients' health and satisfaction. However, most medical practices have not been subjected to such analysis, and even for well-accepted medical practices, the link between process and patient health and satisfaction has often not been established (see ch. 1).

Outcomes of care refer to patient health and satisfaction. In assessments of quality, outcomes acquire importance to the extent that they have resulted from prior medical interventions. But attributing changes to medical care requires distinguishing the effects of care from the effects of the many other factors regarding patients and their environments that also influence health and satisfaction.

Because of these conceptual difficulties, process and outcome measures should be used as complementary indicators of quality rather than alternatives. Process measures acquire validity as indicators of quality only to the extent that they have been found likely to improve or harm patient outcomes. And particular outcomes are valid indicators of quality only to the extent that they can be linked to prior process.

Indicators of the quality of care may be viewed in terms other than their relationship to structure, process, or outcome of care. Indicators may pertain to specific diagnoses, conditions, and procedures or to overall care for a person or episode. Indicators vary in the sources of information required. Evaluating whether appropriate procedures were followed for a certain condition or diagnosis requires examination of patients' medical records, while other indicators, such as a physician's specialty or a hospital's mortality rate, may be published or publicly available. Relevant information may also be drawn from claims to third-party payers, from routinely prepared hospital discharge abstracts, and from special surveys. Indicators may be applied to perform different functions. Some indicators may be used to screen large data bases for cases that are especially likely to entail poor performance. Other indicators may be applied to evaluate care more intensively, perhaps by reviewing the practices documented in medical records.

Aspects of Medical Care To Evaluate

A framework for assessing quality from the perspective of individual consumers starts with the identification of technical and interpersonal aspects of medical care to evaluate. Table 3-1 lists 10 aspects of medical care that surveys of individual consumers (see ch. 2) and the literature have indicated affect the desired outcomes, namely patients' health and satisfaction. A provider's responsiveness to urgent or emergency situations may control whether patients obtain medical care in time for their conditions to be helped. Similarly, referring patients to the appropriate

Table 3-1.—Aspects of Medical Care To Evaluate

1. Responsiveness to urgent/emergency situations
2. Referral to appropriate level of care
3. Humaneness
4. Communication of information
5. Coordination and continuity of care
6. Primary prevention
7. Case finding
8. Evaluation of presenting complaint
9. Diagnosis
10. Management:
    - —Patient education
    - —Referral/consultation
    - —Therapy
    - —Monitoring
    - —Followup

SOURCE: Office of Technology Assessment, 1988

level of care, perhaps through transfer to another facility or referral to a particular specialist, may affect the care that patients receive and the extent to which their medical conditions are improved. How physicians and hospitals respond to people seeking urgent care and handle transfers certainly affects patient satisfaction.

The inclusion of a provider's humaneness and communication of information as aspects of care to evaluate reflects the importance that consumers place on being treated respectfully and on having their conditions and treatments explained to them. People place a high value on physicians' taking the time to answer questions and offer explanations. Although all patients may not want very detailed information, physicians face the difficult task of sensing how much is wanted by a given patient and providing it.

Five of the categories in table 3-1—prevention, case finding, evaluation of presenting complaint, diagnosis, and management—relate to the steps that are taken during an episode of care, regardless of the setting(s) in which care is delivered (see figure 3-1). Having the desired effects on health and patient satisfaction require that patients receive appropriate medical care, both technical and interpersonal, at each of these steps.

Coordination of care is singled out for particular emphasis. Even if each health professional in each setting performed each step appropriately, poor care could result from lack of coordination across professionals, sites, and steps. Researchers have found that continuity improves patient satisfaction and compliance (177), although its importance, like that of other aspects of medical care, varies according to the situation (183).

Possible Indicators of Quality for Individual Consumers

A number of indicators have been suggested for assessing the quality of medical care provided by hospitals and physicians. Tables 3-2 and 3-3 list commonly cited indicators and relate them to the 10 aspects of medical care that are important to consider. Despite the application of many of these indicators in the research literature and the popular press, few have been subjected to rigorous evaluation of their reliability and validity as measures of quality. Moreover, the evaluations that have been performed have found little to support the validity of many commonly used indicators, such as board certification of physicians (477). Nevertheless, possible indicators have been compiled in these tables to illustrate different approaches to measurement and to exemplify the wide range of quality measures that have been suggested or used.

The appropriate indicators for measuring the quality of care depend on the characteristics of the patient and the aspect of quality that is being considered. The indicators in tables 3-2 and 3-3 relate to general characteristics of hospitals and physicians or general review of their patients' cases. If shown to be valid, such indicators could guide a consumer who wished to choose a physician or hospital for all-purpose care. The exception is the volume of specific procedures or diagnoses, such as cardiac bypass surgery, hip replacement, or acute myocardial infarction. People with a condition or others acting on their behalf would probably wish information only on a specific procedure. For other indicators listed, such as physician specialization, evaluation of performance for particular conditions, and hospital mortality rates, the resulting information could relate either to general care or to more specific conditions. Consumers evaluating a particular hospital might wish to know the mix of specialties available or the specialists available to treat one condition. Quality assessors could review medical records across all conditions or restrict the sample to a specific condition. Similarly, hospital mortality rates could pertain to the entire institution, a department, or a procedure or condition.

Tables 3-4 and 3-5 provide selected information on the use of medical specialists and provide a context for understanding how information on specific physicians could help consumers select a physician. As shown in table 3-4, which physician specialists people use depends to a great extent on the patients' age (and sex). The reason is partly that some specialties, such as pediatrics, concentrate on the care of one age group, and partly that most specialties focus their practice on certain medical conditions, which in turn vary according to patient age. Table 3-5 shows for four

Table 3-2.—Possible Indicators of Hospital Quality and Their Relationship to Aspects of Medical Care

Structural indicators:	
Accreditation by Joint Commission on the Accreditation of Healthcare Organizations	(overall performance)
Affiliation with medical school	(overall performance)
Credentialing process to admit physicians to staff	(overall performance)
Medical staff organization	(overall performance)
Ombudsman/mechanism for handling complaints	(overall performance)
Organization of nursing staff	(overall performance)
Proportion of staff graduated from foreign medical schools	(overall performance)
Staff turnover	(overall performance)
Teaching status	(overall performance)
Registered nurses in direct patient care per patient	(overall performance, 4)
Volume of specific procedures or diagnoses	(overall performance, 10)
Scope of services, including emergency facilities and physician services	(overall performance, 1,8,9,10)
Specialization of physicians	(overall performance, 2,5,7,8,9,10)
Procedures of quality assurance committee	(overall performance, 1,2,4,5,6,7,8,9,10)
Active ethics committee	(2,3,10)
Certification of laboratory	(5,7,8)
Structural indicators (cont'd):	
Availability of home health services	(5,10)
Community education program	(6)
Certification of blood bank	(5,10)
Process indicators:	
Disciplinary actions	(overall performance)
Performance for specific medical procedure(s) or condition(s)	(overall performance, 2,4,5,6,7,8,9,10)
Autopsy rates	(8,9,10)
Removal of normal tissue	(8,9,10)
Outcomes:	
Adverse events	(overall performance)
Patient ratings	(overall performance)
Malpractice compensation	(overall performance, 3)
Nosocomial infections	(overall performance, 10)
Hospital mortality rates	(overall performance, 2,5,7,8,9,10)
Measures of functional status	(overall performance, 2,5,8,9,10)
Hospital readmissions	(overall performance, 8,9,10)
Drug and transfusion reactions	(5,8,9,10)

Key to numbers representing aspects of care:
1 = Responsiveness to urgent/emergency situations
2 = Referral to appropriate level of care
3 = Humaneness
4 = Communication of information
5 = Coordination and continuity of care
6 = Primary prevention
7 = Case finding
8 = Evaluation of presenting complaint
9 = Diagnosis
10 = Management

SOURCE: Office of Technology Assessment, 1988.

major age groups the most frequently used physician specialties along with the major causes of hospitalization, disability, and death that they treat.

One might place high priority on assessing the quality of the physician specialties on whom people rely most, namely the primary-care specialties including general or family practice, internal medicine, pediatrics, and obstetrics/gynecology. Or priority might fall to specialties that manage conditions that pose substantial risk to patients, because the conditions require hospitalization or jeopardize mobility or life. People seeking a family physician would benefit from evaluations that spanned the range of conditions a specialty commonly manages, while people choosing a physician for a particular condition would desire information that related to that condition. Whether for overall care or care for specific conditions, the content of a specialty's care could guide quality assessors' selection of cases and outcomes to evaluate.

Known deficiencies in medical care could also guide the choice of what to assess for consumers (186,704). Assessors could focus on the most common or most dangerous hazards to patients or the areas in which errors can be corrected and the greatest benefits for patients achieved.

Table 3-3.—Possible Indicators of Physician Quality and Their Relationship to Aspects of Medical Care

Structural indicators:
Type of medical school (teaching v. nonteaching)
(overall performance)
Trained in medical-school hospital
(overall performance)
Graduate of foreign medical school
(overall performance)
Specialization
(overall performance, 2,5,7,8,9,10)
Volume of specific procedures or diagnoses
(overall performance, 10)
Hospital admitting privileges
(overall performance, 2,5)
Emergency coverage arrangements
(1)

Process indicators:
Disciplinary actions
(overall performance)
Performance for specific procedure or condition
(overall performance, 2,4,5,6,7,8,9,10)
Drug use
(8,9,10)

Outcomes:
Patient rating
(overall performance)
Adverse events
(overall performance)
Malpractice compensation
(overall performance, 3)
Patient drug reaction
(5,8,9,10)

Key to numbers representing aspects of care:
1 = Responsiveness to urgent/emergency situations
2 = Referral to appropriate level of care
3 = Humaneness
4 = Communication of information
5 = Coordination and continuity of care
6 = Primary prevention
7 = Case finding
8 = Evaluation of presenting complaint
9 = Diagnosis
10 = Management

SOURCE: Office of Technology Assessment, 1988.

INDICATORS OF QUALITY SELECTED FOR OTA EVALUATION

Criteria for Selection

In selecting indicators of the quality of medical care for evaluation, OTA considered the perspectives of consumers, the medical profession, research, and policy. As indicated in table 3-6, OTA attempted to incorporate indicators perceived to be valid by consumers and by those in the medical, research, and policy communities. Each of these groups is using certain indicators to assess quality, often without thorough evaluation of the indicators' validity. Subjecting such indicators to intensive examination could validate their appropriateness or elucidate problems with their use.

Since OTA's task is to evaluate indicators of quality that consumers could use to choose physicians and hospitals, the public's requirements for information received high priority. People are most likely to face decisions about medical care for the conditions that have the highest incidence and prevalence in the United States. The most common causes of physician office visits, hospitalizations, disability days, and death were the basis of the entries in tables 3-4 and 3-5. As one would expect, the most frequent afflictions vary by age and sex. In addition, the circumstances and type of medical condition influence how consumers choose providers. One survey organization reported that, on average, 22 percent of consumers selected a hospital on their own, without their physicians' advice; in cases involving accident or injury, however, 33 percent chose the hospital independently. People were also more likely to act on their own in choosing a hospital for general tests and treatment (29 percent) and for illness and maternity (27 percent) than for surgery (17 percent) (320).

Also important in OTA's selection was that the indicators taken together relate to the aspects of care that are important to people (see table 3-1 and ch. 2). People have reported being particularly concerned about humaneness and communication of information, including information on primary prevention (392).

Other considerations in selecting indicators to evaluate hinged on the state of medical knowledge. Given current information and technology, certain events, such as maternal deaths, should

Table 3-4.—Distribution of Office Visits to Physicians, by Physician Specialty and Patient Age, 1985[a]

	Percent of visits by patient age					
Physician specialty	Birth-14 years	15-24 years	25-44 years	45-64 years	≥65 years	Total population
General or family practice	25.0%	35.6%	31.9%	32.0%	29.0%	30.5%
Internal medicine	2.2	6.4	9.1	15.7	22.0	11.6
Pediatrics	55.2	6.0	1.1	0.4	0.2	11.4
Obstetrics/gynecology	0.5	18.8	19.3	4.7	1.4	8.9
Ophthalmology	2.5	4.0	3.9	7.0	13.5	6.3
Orthopedic surgery	2.9	6.2	6.1	6.1	3.4	4.9
General surgery	1.4	4.1	4.5	6.6	6.2	4.7
Dermatology	1.4	6.4	4.6	3.8	3.4	3.8
Psychiatry	0.7	2.3	5.8	3.0	0.9	2.8
Otolaryngology	3.5	2.2	2.3	2.6	2.1	2.5
Urology	0.5	0.8	1.4	2.4	3.5	1.8
Cardiology	0.1	0.3	0.6	3.1	3.8	1.7
Other	4.3	6.7	9.4	12.5	10.5	9.0
Total	100%	100%	100%	100%	100%	100%

[a]Percentages may not add to 100 because of rounding.

SOURCE: U.S. Department of Health and Human Services, Public Health Service, National Center for Health Statistics, unpublished data from the National Ambulatory Medical Care Survey, Hyattsville, MD, Nov. 17, 1986.

occur only rarely, and their occurrence often raises concern about the quality of care. Especially in the past 50 years, medical advances have enabled providers to intervene in the natural progression of many medical conditions, to restore function or to prevent further decline. But most techniques, even well-accepted ones, have not been well evaluated, and many may lack efficacy. Consequently, it is reasonable to restrict evaluations of quality to the application of technologies with demonstrated efficacy and to conditions with efficacious interventions.

By drawing indicators from the different research approaches used to evaluate quality (structure, process, and outcome), OTA hoped to gain insight into advantages and disadvantages of each approach. To ensure the feasibility of its own research, OTA limited its analysis to indicators for which sufficient published and unpublished information existed to support an evaluation.

Reflecting the interest of Congress and other policymakers, OTA paid particular attention to indicators that quality assessors are using or considering, especially for public programs. Also in line with policy interests, OTA wished to target conditions or interventions where quality problems are likely because of overuse or underuse of particular procedures.

Indicators Selected for Evaluation

OTA selected the following eight categories of indicators for intensive evaluation:[2]

- hospital mortality rates, for the overall institution, by department, and by condition or procedure;
- adverse events that affect patients, as exemplified by nosocomial (institutionally acquired) infections in hospitals;
- formal disciplinary actions by State medical boards, sanctions recommended by utilization and quality control professional review organizations (PROs) and imposed by the U.S. Department of Health and Human Services (HHS), and malpractice compensation;
- evaluations of physicians' performance for a specific condition, as exemplified by physicians' care for hypertension;
- volume of services performed in hospitals and by physicians;
- scope of hospital services, with particular emphasis on emergency services, cancer care, and neonatal intensive care units;
- physician specialization; and
- patients' assessments of their care.

---

[2] App. A contains more information about the selection process.

Table 3-5.—Management of Specific Conditions as Possible Indicators of Quality

Patients from birth to 17 years:

- *Pediatrics, general and family practice*
  - General medical exam, including childhood immunizations
  - Earache/otitis media
  - Respiratory symptoms
  - Asthma
  - Anemia
  - Gastrointestinal symptoms
  - Acne
  - Head trauma, including use of skull X-rays
- *Otolaryngology*
  - Otitis media
  - Allergy
- *Orthopedic surgery*
  - Orthopedic impairments
- *Ophthalmology*
  - Vision problems
- *Dermatology*
  - Acne
- *General surgery*
  - Appendectomy
  - Hernia repair

Patients from ages 18 to 44:

- *General and family practice, internal medicine*
  - General medical exam
  - Hypertension (screening and treatment)
  - Respiratory symptoms
  - Allergy
  - Arthritis
  - Pneumonia
- *Obstetrics/gynecology*
  - Prenatal care and delivery
  - Gynecological disorders
  - Complicated pregnancy (including performance of cesarean section)
  - Hypertension (screening and treatment)
- *Orthopedic surgery*
  - Back symptoms/disc disorders
  - Fractures and dislocations
  - Orthopedic impairment
- *Dermatology*
  - Acne
- *Psychiatry*
  - Depression
  - Alcoholism (treatment)
- *General surgery*
  - Hemorrhoids
  - Cholelithiasis
- *Otolaryngology*
  - Hearing impairments

Patients from ages 45 to 64:

- *General and family practice, internal medicine*
  - General medical exam
  - Hypertension (screening and treatment)
  - Diabetes mellitus (screening and treatment)
  - Respiratory symptoms
  - Arthritis
  - Allergy
  - Angina pectoris
  - Pneumonia
  - Influenza
- *Ophthalmology*
  - Vision problems
- *General surgery*
  - Hernia repair
  - Cholelithiasis
  - Malignant neoplasm of the lung
- *Orthopedic surgery*
  - Back symptoms/disc disorders
  - Fractures and dislocations
- *Gynecology*
  - Hypertension (screening and treatment)
  - Diabetes mellitus (screening and treatment)
- *Dermatology*
  - Skin disorders
- *Cardiology*
  - Angina pectoris
- *Otolaryngology*
  - Hearing impairments
- *Urology*
  - Calculus of kidney and ureter

Patients aged 65 and older:

- *General and family practice, internal medicine*
  - General medical exam
  - Hypertension (screening and treatment)
  - Congestive heart failure
  - Ischemic heart disease
  - Diabetes mellitus (screening and treatment)
  - Arthritis
  - Chronic obstructive pulmonary disease
  - Influenza
  - Pneumonia
  - Respiratory symptoms
- *Ophthalmology*
  - Cataract removal
  - Other vision problems
- *General surgery*
  - Cataract removal
  - Malignant neoplasm of lung
  - Malignant neoplasm of breast
  - Varicose veins
- *Cardiology*
  - Congestive heart failure
  - Acute myocardial infarction
  - Ischemic heart disease
- *Urology*
  - Prostatectomy
- *Dermatology*
  - Skin disorders
- *Orthopedic surgery*
  - Fracture of neck of femur
- *Otolaryngology*
  - Hearing impairments

SOURCES: *Morbidity and Mortality Weekly Report*, "Premature Mortality in the United States," 35(2S):1S-11S, Dec. 19, 1986. U.S. Department of Health and Human Services, Public Health Service, National Center for Health Statistics, "1985 Summary: National Hospital Discharge Survey," *NCHS Advance Data*, No. 127, Hyattsville, MD, Sept. 25, 1986. U.S. Department of Health and Human Services, Public Health Service, National Center for Health Statistics, unpublished data from the National Ambulatory Medical Care Survey, Hyattsville, MD, Jan. 16, 1987. U.S. Department of Health and Human Services, Public Health Service, National Center for Health Statistics, unpublished data from the National Health Interview Survey, Hyattsville, MD, Nov. 7, 1986.

Table 3-6.—Considerations in Selecting Indicators of Quality for OTA Evaluation

Consumer interests:
- High-frequency conditions or reasons for seeking care
- Indicators together cover range of what is important to people
- Indicators together relate to general population, particular age-sex categories, and vulnerable groups

Medical interests:
- Conditions for which medical care can alter the natural history
- Events that should not occur
- Conditions or interventions where quality problems are likely from overuse or underuse of particular procedures
- Indicators perceived as valid by medical community

Research interests:
- Information available to support an evaluation
- Indicators that relate to different approaches to assessing quality (structure, process, and outcome)

Policy interests:
- Indicators frequently considered to assess quality
- Indicators being used to assess quality

SOURCE: Office of Technology Assessment, 1988.

Taken together, these eight indicators relate to a range of medical providers, types of medical care, aspects of care, approaches to quality assessment, and sources of data (see table 3-7). Hospital mortality rates and scope of hospital services apply only to hospitals, and physician specialization applies most directly to physicians. Five of the indicators—adverse events, disciplinary actions and malpractice compensation, evaluation of physicians' performance for a specific condition, volume of procedures, and patient ratings—could apply to both physicians and hospitals. This report does not explicitly consider indicators of quality for HMOs and other alternative delivery systems; however, quality assessors could use these indicators to evaluate physicians and hospitals associated with such organized delivery systems as well as physicians and hospitals operating more independently.

All but one of the eight indicators evaluated in this report pertain to the evaluation of general rather than condition-specific care. Only the evaluation of physicians' performance through hypertension screening and management pertains to a specific condition, but the evaluation of other indicators touches on age- and sex-specific conditions for which people frequently seek care. The analysis of hospital mortality rates examines mortality rates for specific departments, such as neonatal intensive care units; and the analysis of volume of procedures examines procedures for several specific conditions, such as appendectomy, hysterectomy, coronary artery bypass graft, total hip replacement, prostatectomy, and acute myocardial infarction. Whether a hospital's scope of services is adequate depends on what medical conditions the hospital treats. Although this report does not explore them in depth, some adverse events, such as maternal death, relate to specific conditions.

Each of the indicators that OTA chose for evaluation is associated with 1 or more of the 10 specific aspects of medical care that were listed in table 3-2. As shown in table 3-7, hospital mortality rates, adverse events, State disciplinary actions, PRO/HHS sanctions, and malpractice compensation could result from deficiencies in any of several aspects of care. Patients' assessments are associated with a number of matters of particular concern to consumers: the responsiveness of a provider to urgent situations, the personal respect or humaneness accorded a patient, the communication of desired information, and the performance of primary preventive activities. Review of the care given for hypertension would give information on almost the entire range of medical care aspects.

The eight indicators encompass the range of approaches to assessing quality: structure, process, and outcome. Two indicators—hospital mortality rates and adverse events that affect patients—enumerate undesirable effects on patient health. Both pertain almost exclusively to physiologic health and physical function. State disciplinary actions, PRO/HHS sanctions, and malpractice compensation are indicators that straddle the process and outcome categories; patients or colleagues may undertake malpractice and disciplinary actions because of providers' negligence in the provision of medical care, but the allegedly negligent behavior may attract notice because of adverse effects on patients' health or satisfaction. The review of physicians' care for a specific medical condition, such as hypertension, entails scrutinizing aspects of the medical care process. Three indicators—volume of procedures provided

Table 3-7.—Issues Addressed by the Indicators Selected for OTA Evaluation

	Hospital mortality rates	Adverse events	State disciplinary actions, PRO/HHS sanctions, and malpractice compensation	Evaluation of physicians' performance: hypertension	Volume of services	Scope of hospital services	Physician specialization	Patients' assessments
Providers:								
Physicians		X	X	X	X		X	X
Hospitals	X	X	X	X	X	X		X
Type of medical care:								
General care	X	X	X			X	X	X
Condition-specific care	X	X		X	X	X	X	X
Aspects of medical care:								
Overall performance	X	X	X		X		X	X
Responsiveness to urgent situations	X	X	X			X		X
Referral to appropriate level	X	X	X	X		X		
Humaneness			X	X				X
Communication of information		X	X	X				X
Coordination and continuity of care	X	X	X	X	X	X		X
Primary prevention		X		X				X
Case finding	X	X	X	X				
Evaluation of presenting complaint	X	X	X	X				X
Diagnosis	X	X	X	X				X
Management	X	X	X	X	X			X
Assessment approach:								
Structure					X	X	X	X
Process		X	X	X				X
Outcome	X	X	X	X				X
Source of data:								
Large data bases	X	X	X		X	X	X	
Chart review		X					X	
Special survey		X						X

SOURCE: Office of Technology Assessment, 1988.

by a hospital or physician during a year, scope of hospital services, and physician specialization —represent structural measures of quality; that is, they all measure the existence of certain medical resources, including expertise and facilities. Patients' assessments of their care occupy a dual position in this schema. Patients' assessments may serve as a measure of patient satisfaction, one of the desired outcomes of medical care. Or patients may rate or report structural and process characteristics of care (e.g., a provider's responsiveness to urgent situations).

Evaluation of the Indicators: General Issues

Applying the method described in appendix C, OTA evaluated the reliability, validity, and feasibility of using each of the eight quality-of-care indicators to inform the public about the quality of physicians or hospitals. *Reliability* relates to whether a measure of the same case will produce the same results on successive trials, *validity* to whether an indicator measures what it purports to measure, and *feasibility* to whether it is practical to use a certain indicator to convey information to the public about quality. Although each indicator raises different considerations, the reader should be alert to certain general issues that relate to many of the indicators and threaten their reliability, validity, and feasibility.

Making reliable comparisons of providers' quality requires that providers be assessed by the same standards and that the measures conform to uniform definitions. But developing information to construct or to interpret each indicator evaluated in this report requires people to make judgments: physicians and other medical professionals to set standards and to review the performance of their peers, judges and public administrators to interpret laws and regulations, statisticians to analyze data, or patients to assess their care. The decisions of experts in a field often differ because the experts have different knowledge and opinions (the problem of interrater reliability). Even the same person may judge the same situation differently at different times (the problem of intrarater reliability). This situation calls into question the reliability of the eventual evaluations of providers' quality. For example, one researcher reported that, among reviewers who had received no training in evaluation, agreement on assessments of medical records approached only 50 percent, no better than chance (479).

Researchers and quality assessors have attempted to mitigate this problem by specifying explicit criteria for reviewers to use. Although this approach may improve interrater reliability, it may simultaneously reduce validity (184). With the use of explicit criteria, reviewers may have little flexibility to take into account what is appropriate for specific patients. In an attempt to realize the advantages and avoid the disadvantages of each method, quality assessors, including PROs, are combining approaches by using patient outcomes or explicit process items to identify problem cases that receive subsequent implicit review (see ch. 5 on adverse events and ch. 7 on evaluations of physicians' performance for a specific condition).

Questions of reliability also arise in connection with common data sources and definitions. Diagnostic information entered on hospital discharge abstracts, a primary source of information for quality assessment, may differ among hospitals because coders use different definitions (166). Even apparently straightforward facts such as death may not be recorded reliably and in any case are subject to differing definitions, depending, for example, on whether the death occurred before or after hospital discharge.

Several considerations threaten the validity of the indicators. As described above, each of the three major approaches to measuring quality—structure, process, and outcome—has shortcomings. Structural measures describe the potential of a hospital or physician to deliver good quality care, but cannot guarantee it. Structure is at best a necessary, but not a sufficient, condition for good quality care. Elements of the medical care process have validity as predictors of the quality of care only to the extent that research has established their efficacy in achieving desired patient outcomes. Conversely, to establish the validity of an outcome measure, one must be able to attribute the results to prior medical care, as opposed to the host of other factors that may influence what happens to patients.

Regardless of the approach, quality assessors face the problem of how to set the criteria and standards by which to evaluate medical providers. Following the work of Donabedian, criteria refer to the elements to be measured in an evaluation, and standards pertain to what is considered acceptable or good (184). The validity of the criteria and standards that are set is threatened by dependence on the judgments of experts. Some problems arise because of the subjectivity of experts' decisions about what does or does not constitute good quality care. But a perhaps more serious problem is the lack of scientific information on the efficacy and safety of most medical practices. The less information comes from studies documenting efficacy and safety, the greater the role of experts' judgments, with all their subjectivity.

An additional validity issue concerns the generalizability of results and whether evaluations should relate to a provider's entire practice or only to specific conditions. Each level of aggregation has a role to play in quality assessment and complements the other. How a physician or hospital manages a specific condition, such as hypertension or coronary artery bypass surgery, has clinical relevance to other health professionals and to individuals or organizations seeking a provider for a certain purpose. As a rule, however, one cannot generalize from how well a medical provider handles one condition to how well that provider handles other conditions and performs overall. Conversely, evaluations across the range of conditions that a medical provider usually manages would convey information to quality reviewers about the provider's overall performance and could help people seeking a primary care physician or a physician in a certain specialty.

In the area of feasibility, inadequate data pose the most important and most pervasive problem. Both outcome and process measures of quality require clinical data that are generally lacking in routinely available data bases, such as providers' insurance claims and hospital discharge abstracts. Furthermore, because existing sources do not combine ambulatory and inpatient records, reviewers are unable to evaluate an episode of care and attribute responsibility for the results among providers.

The underlying question that remains is whether any of the possible indicators of medical care quality provide reliable and valid assessments that consumers can use to select physicians and hospitals. The subsequent chapters of this OTA report address that question for the eight selected indicators.

Chapter 4

Hospital Mortality Rates

CONTENTS

Page

Introduction . . . 71
Reliability of the Indicator . . . 72
Validity of the Indicator . . . 81
  Intelligibility of Hospital Mortality Rates as an Indicator of Quality . . . 81
  When To Measure Death . . . 82
  Adjusting for Patients' Risk of Dying . . . 82
  Validation of Hospital Mortality Rates Against the Process of Medical Care . . . 85
  Comparisons of Hospital Mortality Rates With Other Potential Measures of Quality . . . 86
  Comparison of Different Hospital Mortality Analyses . . . 86
  Longitudinal Analyses . . . 91
Feasibility of Using the Indicator . . . 93
  Construction of the Indicator . . . 93
  Intentional Manipulation of the Indicator . . . 93
  Dissemination of Information About the Indicator . . . 94
Conclusions and Policy Implications . . . 97

Box

Box *Page*

4-A. Selected Sources of Information About Hospital-Specific Mortality Rates . . 95

Figure

Figure *Page*

4-1. Scoring of Patients Under the APACHE II System for Classifying Severity of Disease . . . 84

Tables

Table *Page*

4-1. Comparison of HCFA's 1984 and 1986 Hospital Mortality Analyses . . . 72
4-2. Characteristics of Hospital Mortality Studies Reviewed by OTA . . . 73
4-3. Results of Hospital Mortality Studies Reviewed by OTA . . . 74
4-4. Comparison of Hospital Mortality With Structural and Other Outcome Indicators in Hospital Mortality Analyses Reviewed by OTA . . . 87
4-5. Comparison of Hospital Mortality Rates: Arizona . . . 91
4-6. Number of Hospitals Found To Be High-Mortality Outliers by HCFA in 1984 and 1986, Selected States . . . 92

Chapter 4

Hospital Mortality Rates

INTRODUCTION

Differences in patient death rates seem on their face a valid way to distinguish good quality health care providers from poor quality providers; death is an outcome that is almost always bad,[1] and medical practice is devoted, at least in part, to postponing death. Differential mortality, or survival, has long been used as a measure of efficacy in health care technology assessments and as an indicator, albeit crude, of the health status of particular populations. Medical encounters can be dangerous (318,555,595), adding to the possibility of death from a hospital encounter.

Almost half the deaths in the United States every year occur in hospitals, although only about 3 percent of hospital admissions end in death (667). Although many deaths in hospitals occur because nothing more could be done for the patients involved, a substantial portion of the deaths are believed to be avoidable. Hospital-related mortality can result from various factors that are subject to control, including poor infection control, inadequate or inappropriate use of medication, falls as a result of poor supervision, mistakes during surgery, and inappropriate discharge.

Although the use of patient death rates to compare the quality of care delivered by specific health care providers has been expanding, it has also been controversial. The major problems with the use of hospital mortality rates as a quality indicator are that mortality can result from many factors other than poor quality care and that techniques to adjust for such factors are generally inadequate. In addition, there are theoretical and practical issues regarding the appropriate period of time for an analysis. Over what period of time is a death to be defined as related to hospital care? Another issue regarding time is the period covered in the analysis. Most releases of information on hospital mortality rates have included data for a single year, but critics argue that data over a longer period of time may be needed, given the uncertainties about the indicator. Yet another significant issue is the level of aggregation of hospital mortality rates. Should rates be aggregated across the hospital as a whole? If not, at what level of diagnostic coding should the data be totaled? Finally, it is important to validate hospital mortality rates against criteria related to the process of care; this validation is only beginning.

Perhaps the most visible and controversial releases of hospital mortality data have been the 1984 and 1986 analyses of the Health Care Financing Administration (HCFA), which is part of the U.S. Department of Health and Human Services (640,647). The HCFA releases illustrate well the critical issues surrounding the use of hospital mortality rates as an indicator of the quality of care. Both analyses were conducted with data derived from hospital claims filed for the purpose of Medicare reimbursement, although the 1986 analysis added information about deaths derived from Social Security Administration files (see table 4-1 for a summary of differences between the 1984 and 1986 HCFA analyses). The 1986 analysis differed in level of analysis, in the way conditions and procedures were aggregated, in the period of time after hospital admission during which hospitals were counted, in calculation methods, and in the type of information released.

A number of other analyses of hospital mortality data have been conducted along the same basic lines as the HCFA analyses, that is, using data from hospital discharge abstracts to adjust for patients' risk of dying (80,81,189,448,462,526); other analyses have adjusted for patients' risk of dying using clinical data (190,352,353,588,589, 590) as well as proxies such as age. Few have attempted to validate statistical results against a process criterion (190,279,353,462).

OTA reviewed in depth studies whose purpose was to develop a valid technique to adjust hospital mortality statistics for patients' risk of dying. Not included were studies whose primary purpose was to test the validity of structural measures of

---

[1] It has been argued that in some cases death would be preferable to life; definitions of life and death are not as simple as they once seemed (632).

Table 4-1.—Comparison of HCFA's 1984 and 1986 Hospital Mortality Analyses

	HCFA's 1984 analysis[a]	HCFA's 1986 analysis[b]
Data base	Claims filed for Medicare reimbursement[c]	Claims filed for Medicare reimbursement and information from the Social Security Administration about date of death
Hospital population	Short-term acute care hospitals (some hospices included inadvertently)	Short-term acute care hospitals (some hospices included inadvertently)
Patient population	All Medicare patients, both aged and disabled	All Medicare patients, both aged and disabled
Period of time during which deaths were counted	In-hospital deaths	Death within 30 days of last hospital admission
Hospital "risk group"[d]	All discharges	Last admission
Measures used to adjust for patients' risk of dying	Average age of Medicare patients; proportion male; proportion black; proportion neither black nor white; State average length of stay; 50 most frequent diagnosis-related groups (DRGs); all cancer DRGs; 30 DRGs associated with most frequent DRGs; weighted by number of Medicare discharges	Age group; sex; comorbidities tailored to diagnostic group; prior hospital admissions in the year preceding death; whether patient was transferred from another hospital
Level of analysis	Hospital	Patient, then hospital
Levels of aggregation	Hospital overall, and 9 DRG categories	Hospital overall, and 17 diagnostic risk groups
Calculation method	Multiple linear regression	Logistic regression
Information released	Outlier hospitals only	All hospitals, with actual and expected mortality rates for each category

[a] U.S. Department of Health and Human Services, Health Care Financing Administration, "Medicare Hospital Mortality Information 1984," Washington, DC, Mar. 10, 1986.
[b] U.S. Department of Health and Human Services, Health Care Financing Administration, *Medicare Hospital Mortality Information 1986* (Washington, DC: U.S. Government Printing Office, 1987).
[c] Assembled in HCFA's MEDPAR file.
[d] Denominator.

SOURCE: Office of Technology Assessment, 1988.

quality against hospital mortality as a standard. In addition, the OTA review included releases of crude mortality rates (55,115 116,478) to compare their results with the rates adjusted in various ways. All studies were reviewed using the procedure and checklist described in appendix C.[2] Table 4-2 lists the studies reviewed by OTA, and indicates when they were conducted, the sources of data used, the patient and hospital types that were included, and the years in which data were collected. Table 4-3 shows the diagnoses and procedures included in the analysis, when death was measured, the adjustments for patients' risk of dying,[1] the level of analysis, and the results of each study.

The remainder of this chapter consists of an evaluation of the reliability, validity, and feasibility of using hospital mortality rates as an indicator. Conclusions and policy implications are outlined in the final section of the chapter.

RELIABILITY OF THE INDICATOR

Whether hospital mortality rates are a valid indicator of the quality of care depends on the reliability of the data on which analyses of mortality rates are performed and the reliability of the data against which results of analyses are validated. Some aspects of the data base for hospital mortality analyses have been of longstanding concern (166,167). There is reason to believe that hospital data sources vary widely in completion and accuracy; rarely have hospital mortality analy-

---

[2] The way studies were selected for review and descriptions of the individual studies can be found in OTA's technical working paper, "Hospital Mortality Rates as a Quality Indicator" (187).

Table 4-2.—Characteristics of Hospital Mortality Studies Reviewed by OTA

Study[a]	Patient population	Source of data	Hospital types included or excluded	Years data collected	Sample size
Bunker, et al., 1969 (108)	All	Hospital medical records	Included military, National Institutes of Health, teaching and community general hospitals (all volunteers)	(1959-62) 4 years	34 hospitals; 856,000 patients; 16,840 deaths
Moses and Mosteller, 1968 (441)	All	Same as Bunker, et al., 1969	Same as Bunker, et al., 1969	Same as Bunker, et al., 1969	34 hospitals; 141,914 patients; 1,844 deaths
Roemer, et al., 1968 (526)	All nonobstetric	State of California hospital annual reports	Hospitals in Los Angeles County, including Veterans Administration and municipal[b]	1964	33 hospitals
Goss and Reed, 1974 (259)	All nonobstetric	Deaths: death certificates	102 short-term general hospitals in New York City	1971	50,000 deaths
Stanford Center for Health Care Research, 1974 (588), 1976 (589):					
Extensive Study	All	Commission on Professional and Hospital Activities' Professional Activities Study	Short-term hospitals	1972	1,244 hospitals; 558,856 patients
Intensive Study	All	Same as used in Extensive Study, plus data collected at hospital sites	Short-term hospitals randomly selected from a sample stratified by size, teaching status, cost per patient day, and a crude estimate or surgical mortality	May 1973-Feb. 1974 (9 months)	17 hospitals; 8,593 patients
NAS, 1977 (448)	Males	Veterans Administration Patient Treatment File	Veterans Administration hospitals, including psychiatric hospitals	1970-75 (6 years)	More than 200,000 surgeries
Knaus, et al., 1986 (353)	Adults only; no coronary artery bypass graft	Hospital and medical records and questionnaire data	Hospitals volunteering to be in the study	Average of 5 months[c]	13 hospitals; 236 patients
U.S. DHHS, HCFA, 1986 (640)	Medicare patients, all ages	Medicare billing file	Short-term general hospitals	1984	Not given
Blumberg, 1987 (80), 1988 (81)	All	Maryland Health Services Cost Review Commission data base (based on discharge abstracts)	All Maryland hospitals except 10	April 1984-March 1985 (1 year)	45 hospitals, 8,745 cases
New York State Department of Health, 1987 (462)	All	New York State Department of Health Statewide Planning and Research Cooperative System	Excluded children's hospitals, one maternity hospital, a cancer hospital, several rehabilitation hospitals, and an eye-ear-throat hospital	1984	Not given
Rust, et al., 1987 (545)	Newborns	Birth and death certificates, State of California	NA[d]	1980-84	340 hospitals; 2.5 million babies
Dubois, et al., 1987 (189,190)	All	Modified version of the Uniform Hospital Discharge Data Set, aggregated to the hospital level	American Medical International, Inc. hospitals selected to be geographically representative[e]	Six-month period in 1985	93 hospitals; 205,000 hospital discharges
U.S. DHHS, HCFA, 1987 (647)	Medicare patients, all ages	Medicare billing data base (MEDPAR); Social Security Administration records (for deaths)	Short-term general hospitals	1986	10 million admissions
DesHarnais, et al., 1988 (173)	a. All except newborns, transfers to other short-stay hospitals, stays of less than 1 day	a. Commission on Professional and Hospital Activities data base		a. 1983-84	a. 300 hospitals
	b. Medicare patients, all ages	b. Medicare billing data base (MEDPAR)	Short-term general hospitals	b. 1984	b. Not given

[a] Studies are listed in chronological order. Numbers in parentheses refer to numbered entries in the reference list at the end of this report.
[b] The hospitals were chosen to represent range of medical staff organization types (loosely to highly structured).
[c] Data were collected either on consecutive patients or on every second or third patient until a specified number of patients was reached.
[d] NA = Not applicable.
[e] Hospitals were nonteaching, nongovernmental, and proprietary.

SOURCE: Office of Technology Assessment, 1988.

Table 4-3.—Results of Hospital Mortality Studies Reviewed by OTA

Study[a]	Diagnoses and/or procedures included	Dependent variable	Adjustments	Level of analysis	Results: percent of variance in crude mortality explained (R^2), if available	Relation to validation standard for process of care
Bunker, et al., 1969 (108)	a. All operations requiring general anesthesia	a. Inhospital death within 6 weeks following surgery	a. Operation	a. Patient	a. Standardization for operation reduced variation in death rates from 25:1 (for crude rate) to 9:1	a. —[b]
	b. Sixty-two "middle death rate" surgeries requiring general anesthesia	b. Same as "a"	b. Age, operation, physical status[c], shock-likelihood index[d], sex, year of operation, any previous operations during same or previous month (all separately), and a multiple adjustment for age, physical status, and operation	b. Patient	b. When age, operation, and physical status were combined, the ratio for the most extreme procedure was reduced to 3:1	b. —
Moses and Mosteller, 1968 (441)	Six specific procedures combined (gallbladder surgery, gastric resection, hysterectomy, cystoscopy, open reduction of femur, thoracic or lumbar laminectomy)	Inhospital death within 6 weeks following surgery	Age, operation, physical status, frequency of all six procedures combined	Patient	10:1 ratio between hospital with highest adjusted mortality rate and hospital with lowest adjusted mortality rate for six procedures	—
Roemer, et al., 1968 (526)	Nonobstetric admissions	Inhospital death	Hospital average length stay corrected for occupancy rate	Hospital	$R^2 = .59$	—
Goss and Reed, 1974 (259)	Nonobstetric admissions	Inhospital death	Hospital average length of stay corrected for occupancy rate	Hospital	$R^2 = .35$	—
Stanford Center for Health Care Research, 1974 (588), 1976 (589)						
Extensive Study	a. Nine surgical categories and 2 medical categories combined	a. Inhospital death	a. Age, sex, admission findings[f], height to weight index, severity scores (stage of disease) for principal procedures or diagnoses; comorbidities, and additional procedures, staged for severity (tailored to procedure or diagnosis)	a. Patient	a. $R^2 = .30$[e]	a. —
	b. 1. Intra-abdominal artery operations 2. Selected biliary tract surgery 3. Hip fracture diagnosis with no other trauma 4. Amputation of lower limb with no current trauma 5. Large bowel operations with cancer diagnosis 6. Biliary tract diagnosis-nonsurgical 7. Ulcer diagnosis-nonsurgical 8. Arthroplasty of the hip 9. Stomach operations-ulcer 10. Large bowel operations-specified diagnoses	b. Inhospital death	b. Same as "a"	b. Patient	b. 1. $R^2 = .241$ 2. $R^2 = .051$ 3. $R^2 = .089$ 4. $R^2 = .078$ 5. $R^2 = .060$ 6. $R^2 = .079$ 7. $R^2 = .110$ ISMRs were significantly different among hospitals in surgical categories 1-5 and nonsurgical categories 6 and 7 8. $R^2 = .043$ 9. $R^2 = .112$ 10. $R^2 = .087$ 11. $R^2 = .122$ 12. $R^2 = .194$ 13. $R^2 = .107$ 14. $R^2 = .125$	

Table 4-3.—Results of Hospital Mortality Studies Reviewed by OTA—Continued

Study[a]	Diagnoses and/or procedures included	Dependent variable	Adjustments	Level of analysis	Results: percent of variance in crude mortality explained (R²), if available	Relation to validation standard for process of care
	11. Hip fracture diagnosis with other trauma 12. Amputation of lower limb with current trauma 13. Stomach operations with cancer diagnosis 14. Stomach operations with other diagnosis 15. Large bowel operations with other diagnosis 16. Fracture of the pelvis 17. Fracture of the shaft of the femur				15. $R^2 = .201$ 16. $R^2 = .081$ 17. $R^2 = .156$ ISMRs were not significantly different among hospitals in surgical categories 8-17	
Intensive Study	Fifteen surgical categories combined	Death within 40 days of surgery, death with 40 days of surgery or severe or moderate morbidity at 7 days after surgery	Age; sex; extent of insurance coverage; level of stress; anesthetist's rating of physical status; cardiovascular status; whether surgery was emergency; extent of surgical disease; quadratic function combining age and physical status	Patient	Logistic regression not significant for death alone[h]; Linear regression found significant main effect for hospital[i,j] Adjusted "death or morbidity"[k] ratio ranged from 0.55 to 1.41, a threefold difference	—[g]
NAS, 1977 (448)	a. Three major surgical specialties taken together (general, urologic, orthopedic)	a. Inhospital death within 40 days of surgery	a. Age; other diagnoses; emergency or not; other surgeries before or during selected operation	a. Patient	a. Indirectly standardized mortality rates varied from 0.3 to 1.6 percent. Six of 131 hospitals fell outside a range of 2 standard deviations above and below the mean	a. —[l]
	b. All surgeries performed in VA psychiatric hospitals	b. —	b. None	b. Hospital	b. Crude mortality rate was 7.1% compared to 3.3% in VA general hospitals	b. —
Knaus, et al., 1986 (353)	Patients in intensive care units	Inhospital death	APACHE II (acute physiologic scores; age points; chronic health points; emergency surgery status)	Patient	$R^2 = .99$	Number of daily therapeutic intervention points was significantly higher at hospital with lowest adjustment mortality rate[m,n]
U.S. DHHS, HCFA, 1986 (640)	a. All[o]	a. Inhospital death	a. Average age of Medicare patients; proportion male; proportion black; proportion neither black nor white; State average length of stay; 50 most frequent DRGs; all cancer DRGs; 30 DRGs associated with 50 most frequent DRGs; weighted by number of Medicare discharges	a. Hospital	a. $R^2 = .59$[p]	a. —

Table 4-3.—Results of Hospital Mortality Studies Reviewed by OTA—Continued

Study[a]	Diagnoses and/or procedures included	Dependent variable	Adjustments	Level of analysis	Results: percent of variance in crude mortality explained (R^2), if available	Relation to validation standard for process of care
	b. Nine DRG categories: 1. Pneumonia (DRGs 089-090) 2. Coronary artery bypass surgery (DRGs 106-107) 3. Pacemaker implant (DRGs 115-116) 4. Acute myocardial infarction (DRGs 121-123) 5. Congestive heart failure (DRG 127) 6. Gastrointestinal bleeding (DRGs 174-175) 7. Major joint surgery (DRG 209) 8. Transurethral prostatectomy (DRGs 336-337)	b. Innospital death	b. Average age of Medicare patients, race, sex (all at the hospital level of aggregation)	b. Hospital	b. 1. $R^2 = .053$ 2. $R^2 = .007$ 3. $R^2 = .003$ 4. $R^2 = .019$ 5. $R^2 = .020$ 6. $R^2 = .005$ 7. $R^2 = .068$ 8. $R^2 = .009$	b. —
Blumberg, 1987 (80), 1988 (81)	a. High-risk surgeries	a. Inhospital death	a. Age, sex, type of admission (urgent, emergency), source of admission, risk level of procedure, risk level of comorbidities	a. Patient	a. One of more than 41 hospitals had death rates "deserving review" but not statistically significant. Two other hospitals had lower than expected death rates bordering on significance	a. —
	b. Trauma v. nontrauma and the following surgical categories: Nervous system Respiratory Cardiovascular Gastrointestinal Urinary Musculoskeletal	b. Inhospital death	b. Same as "a"		b. Little variation in trauma cases; substantial variation in nontrauma, gastrointestinal and cardiovascular categories ($Chi^2 = 4$ or more)	b. —
New York State Department of Health, 1987 (462)	a. All medical/surgical, all ages	a. Inhospital death	a. Average age, proportion males; proportion black; proportion neither black nor white; case-mix severity;[r] severity surrogates[s]	a. Hospital	a. $R^2 = .86$[q]	a. Overall, 3% of cases were found to have quality problems
	b. All Medicare discharges		b. Same as above	b. Hospital	b. $R^2 = .781$[t]	b. See "a"
	c. All medical/surgical, under 65		c. Same as above	c. Hospital	c. $R^2 = .92$[u]	c. See "a"
	d. Obstetrics-nursery		d. Same as "a" plus Medicaid as source of payment	d. —	d. $R^2 = .37$[v]	d. —
Rust et al., 1987 (545)	Perinatal (fetal and neonatal)	Death of fetus of 20 weeks or more gestation; death within 28 days of birth	Birthweight, sex, race, multiple births	Patient	$R^2 = .80$	—
Dubois, et al., 1987 (189,190)	a. All	a. Inhospital death	a. Age (percent older than 70); percent admitted from emergency department; percent admitted from nursing home; case-mix index based on DRG weights; average length of	a. Hospital	a. $R^2 = .64$[w]	a. See "b"

Table 4-3.—Results of Hospital Mortality Studies Reviewed by OTA—Continued

Study[a]	Diagnoses and/or procedures included	Dependent variable	Adjustments	Level of analysis	Results: percent of variance in crude mortality explained (R^2), if available	Relation to validation standard for process of care
			stay; hospital size; percentage of Medicaid patients; occupancy rate; discharges to nursing home			
	b. 1. Myocardial infarction (DRGs 121,122,123) 2. Pneumonia (DRGs 79 and 89) 3. Cerebrovascular accident (DRG 14)		b. 1. APACHE II score; body-system score[x] 2. APACHE II score; body-system score 3. APACHE II score; body-system score; presence or absence of a mass effect or blood on head CT scan	b. Patient	b. Actual death rates were within 95% confidence intervals for predicted death rates for all 3 conditions	b. Explicit review—no differences between high and low outliers. Implicit review—after adjustment for patients' risk of dying and for oversampling. Findings were as follows between high- and low-mortality hospitals in estimated rate of preventable deaths in total patient population: 1. No significant difference 2. Significant difference 3. Significant difference
U.S. DHHS, HCFA, 1987 (647)	a. All[0]	a. Death within 30 days of last admission	a. Age group; sex; comorbidities tailored to diagnostic group[aa]; prior hospital admissions in the year preceding death; whether patient was transferred from another hospital	a. Patient[y]	a. $R^2 = .59$[z]	a. —
	b. Specific diagnostic categories High risk: 1. Cancer 2. Stroke 3. Severe acute heart disease 4. Severe chronic heart disease 5. Gastrointestinal catastrophes 6. Metabolic and electrolyte disorders 7. Pulmonary disease 8. Renal disease 9. Sepsis 10. Severe trauma Low risk:	b. Same as "a"			1. $R^2 = .37$[bb] 2. $R^2 = .08$ 3. $R^2 = .12$ 4. $R^2 = .07$ 5. $R^2 = .12$ 6. $R^2 = .22$ 7. $R^2 = .22$ 8. $R^2 = .26$ 9. $R^2 = .09$ 10. $R^2 = .10$ 11. $R^2 = .02$ 12. $R^2 = .04$ 13. $R^2 = .10$ 14. $R^2 = .14$ 15. $R^2 = .26$ 16. $R^2 = .10$	

Table 4-3.—Results of Hospital Mortality Studies Reviewed by OTA—Continued

Study[a]	Diagnoses and/or procedures included	Dependent variable	Adjustments	Level of analysis	Results: percent of variance in crude mortality explained (R^2), if available	Relation to validation standard for process of care
	11. Ophthalmologic disease 12. Gynecologic disease 13. Low risk heart disease 14. Gastrointestinal disease 15. Urologic disease 16. Orthopedic conditions					
DesHarnais, et al., 1988 (173)	a. All except newborns (CPHA data base)	a. Inhospital death	a. (1) Age group (0-64, 65-74, 75 +), presence of comorbidities modeled separately for each DRG cluster[cc]	a. Patient	a. R^2 = 81 (1983 data) R^2 = 84 (1984 data)	a. —
			(2) Age; sex; race; existence of secondary diagnoses; cancer except skin cancer as a secondary diagnosis; risk of death associated with principal diagnosis; risk of death associated with first Class I operative procedure; risk associated with comorbidity having the highest risk; number of secondary diagnoses (except complications) where the risk of death was greater for the secondary diagnosis than for the DRG cluster itself			
	b. All (HCFA data base)	b. Inhospital death	b. Same as "a"	b. Same as "a"	b. R^2 = 48	b. —

Abbreviations: ALC = Alternative Care; CPHA = Commission on Professional and Hospital Activities; DRG = diagnosis-related group; HCFA = Health Care Financing Administration; ISMR = Indirectly Standardized Mortality Ratio.

[a] Numbers in parentheses refer to numbered entries in the reference list at the end of this report.

[b] Dash (—) indicates no attempt was made to validate results against process of care.

[c] Anesthetists' ratings.

[d] Combination of age and physical status.

[e] A.B. Flood, Associate Professor, Medical Humanities and Social Sciences Program, College of Medicine, University of Illinois, Urbana, IL, personal communication, Sept. 17, 1987.

[f] For blood pressure, temperature, hemoglobin, hematocrit, urine sugar, and albumin.

[g] Another study, the Service Intensive Study (SIS), examined the variation in clinical services received and outcomes achieved by all (N = 603,580) patients discharged from 17 IS hospitals during the study period 1970-73 (214,221). Thus, the SIS differed from the IS by: including 3 years of patient outcomes; excluding interview and other obtrusively (relative to the ES) collected IS data; and including data for all patients, not just the surgical patients whose care was emphasized in the ES and IS. The SIS found that lower death rates were significantly related to the receipt of more intensive services, and that higher death rates were related to the duration of services (that is, the number of days in the hospital).

[h] For the outcome "death within 40 days of surgery or severe morbidity on the seventh postoperative day" results were a 10:1 difference between the highest and lowest mortality hospitals (0.37 to 3.7 percent) before a Bayesian adjustment, and 3:1 after a Bayesian adjustment.

Table 4-3.—Results of Hospital Mortality Studies Reviewed by OTA—Continued

[i] And significant interaction between hospital and difficulty of procedure.

[j] Depended on surgical category, but generally, age, physical status, stage of disease, and quadratic function were significant.

[k] Outcome was death within 40 days of surgery or moderate or severe morbidity at 7 days after surgery.

[l] Process of care evaluations were done for a subset of hospitals and patients (12 general hospitals, 595 cases), but the results were not compared to the hospital mortality results. Process of care criteria included the fraction of surgical patients given selected initial examinations, given specific patient education, and given home-care instruction or a follow-up appointment.

[m] Knaus and his colleagues found that the major portion of increased therapy given at Hospital 1 (the hospital with the lowest mortality rate) came from frequent laboratory testing, dressing changes, and chest physiotherapy, which resulted from extensive reliance on a clinical protocol, and not from increased use of unique technologies such as ventilators or pulmonary artery catheters.

[n] Knaus and colleagues also consider interaction and coordination of staff to be process measures, but OTA considers them structural measures.

[o] Note: Medicare patients only. Note further that all Medicare included Medicare patients of all ages, not just those 65 and over.

[p] State average length of stay explained most of the variation in mortality. Age, sex, and race variables were not significant.

[q] Fifteen variables were significant, including proportion of transfers from long term care, average age, percent discharged to other hospitals, percent with residence in same county as hospital, percent with length of stay longer than 90 days, percent with ALC days, and case-mix measure (278).

[r] Each of 50 DRGs with highest number of *deaths* (as opposed to admissions, as used in the 1984 HCFA analysis [640]); each DRG with the same diagnosis as those 50 DRGs; all remaining cancer DRGs; each hospitals's predicted mortality rate based on Statewide rate for DRG.

[s] Proportions of: unscheduled admissions, discharges to another acute care facility, transfers from a hospital, discharges from alternate care, discharges from same county as the hospital, number of transfers from a hospital less number of discharges to a hospital divided by total number of discharges ("net migration"), percent of patients with length of stay greater than 90 days.

[t] Nineteen variables were significant, including proportion black, percent of transfers from residential health care facilities, and case-mix index.

[u] Seventeen variables were significant, including percent black, percent transfers from other hospitals, proportion with residence in same county as hospital, proportion with length of stay greater than 90 days excluding ALC stay, proportion with ALC days, and case-mix.

[v] Proportion males, proportion with Medicaid as primary or secondary payor, and proportion with length of stay greater than 90 days excluding ALC stay are all significant.

[w] Four variables were significant: Age (percent older than 70); percent admitted from emergency department; percent admitted from nursing home; case-mix index based on DRG weights.

[x] The body-system score was a comorbidity scale for each patient that reflected the number of body systems (e.g., cardiovascular) that were affected by any of 50 comorbidities present on the day of admission.

[y] Note that analysis was done at the patient level and then aggregated to the hospital level.

[z] Almost all of the variance was explained by 10 variables: age over 85, severe acute heart disease (as a case mix variable), sepsis, pulmonary disease, cancer as a comorbidity, cerebrovascular accidents, renal disease as a comorbidity, metabolic and electrolyte disturbances, severe chronic heart disease, and age between 70 and 74.

[aa] Any of four additional diagnoses (of cancer, chronic liver disease, chronic renal disease, chronic cardiovascular disease, chronic pulmonary disease, cerebrovascular degeneration/chronic psychosis, hypertensive disease, or diabetes) beyond the principal diagnosis.

[bb] H. Krakauer, Office of Medical Review, Health Standards and Quality Bureau, Health Care Financing Administration, U.S. Department of Health and Human Services, personal communication, Baltimore, MD, Mar. 7, 1988.

[cc] Comorbidities were based on ICD-9-CM codes. Codes that were clearly complications were not considered comorbidities.

SOURCE: Office of Technology Assessment, 1988.

ses reported checking carefully the reliability of data sources. Reliability is of particular concern for hospital mortality analyses. As currently constructed, such analyses are based on small numbers and data for single years. Differences in coding, interpretation, and aggregation across time, across coders or reviewers, and across hospitals could substantially affect hospital comparisons.

Evidence indicates that errors in diagnostic labeling are fairly common (166,167,614). These findings are not surprising given the amount of subjectivity that still exists in coding (77). Errors can be made by the physicians who diagnose the patients' condition and by medical records personnel who transform the diagnoses into universal codes, such as those used in the International Classification of Diseases (ICD-9 codes) and those used for diagnosis-related groups (DRG codes). Random errors in diagnostic labeling undoubtedly exist and generally are not of concern when comparing mortality rates across hospitals, but systematic errors in diagnostic labeling could affect the comparisons. For example, a hospital would have an artificially low expected rate of death from pneumonia if it included in the diagnostic category for pneumonia patients who actually had a less serious illness, such as bronchitis (190). The relationship between tendencies to have coding errors and quality-of-care problems, however, remains unclear.

Using data reported to HCFA by hospitals seeking Medicare reimbursement, the HHS Office of the Inspector General (OIG) found a 20.3 percent error rate in coding across hospitals (304,660). The study was conducted with data from October 1984 to March 1985. A significant number of the errors favored the hospitals; that is, the hospitals were paid more for the hospital stay than they would have been if the correct codes had been submitted (so-called DRG "creep"). A common error was the transposing of principal with secondary diagnoses. A statistically nonsignificant trend was found for differences by hospital bed-size, with smaller hospitals tending to upgrade patient diagnoses. Potentially, this upgrading could lower small hospitals' adjusted mortality rates.[3] Bed-size was the only hospital characteristic used in the analysis. In another arm of the study, the OIG found a higher incidence of DRG "creep" in cases that were discharged prematurely (660a). Hospitals commenting on the 1986 HCFA analysis also reported miscoding of diagnoses so that secondary diagnoses were recorded as principal diagnoses, and vice versa.

In a study that used data from non-Medicare as well as Medicare patients, Dubois and colleagues found a rate of coding errors across hospitals similar to that found by the OIG study (20 percent); but they found that the error rate did not differ significantly between high- and low-outlier[4] hospitals (190). Thus, in this study, coding errors seemed not to be responsible for differences in hospital mortality rates.

Another potential source of differences among hospitals, and thus unreliability in the data, is the extent to which secondary diagnoses are recorded. Consistent recording of secondary diagnoses is essential when such diagnoses are used to indicate comorbidities, a commonly used source of information about patients' risk of dying (172,353, 588,589,590,640,647). In connection with an analysis of hospital mortality rates, the Commission on Professional and Hospital Activities found substantial variation among hospitals in the extent to which they recorded secondary diagnoses (172). When secondary diagnoses are used as proxies for comorbidities, lack of documentation could affect a hospital's expected mortality rate.

The reliability of information about the patient's clinical status on admission can be affected by incomplete entries or inconsistency across raters in recording the information that is available. Incomplete coding of clinical data is a major drawback to the use of patient classification systems based largely on clinical data (94,352, 353). When the State of Pennsylvania decided to publish outcome statistics adjusted with clinical data, for example, it simultaneously implemented a requirement that all hospitals use the same classification system, so that the needed data would be available from all hospitals (41,427). Presumably such a requirement would encourage more consistent recording of such data.

---

[3] The Inspector General's study did not include a review of mortality rates.

[4] After adjusting for patients' risk of dying, analyses estimate for each hospital an expected mortality rate. They then compare the hospital's actual mortality rate to the expected one. Typically, hospitals whose actual rates exceed the expected rates by more than 1.96 standard errors are considered high outliers, and hospitals whose actual rates fall below the expected by more than 1.96 standard errors are considered low outliers. This type of analysis assumes that hospital mortality rates follow a normal distribution, although that assumption has not been validated. See Blumberg and DesHarnais for further discussion of statistical issues surrounding hospital mortality analyses (77,172). In addition, the General Accounting Office is preparing a report on Medicare's use of patient outcome data (626).

Interrater reliability for the clinically based patient classification systems that are being used in mortality analyses (94,352,353) seems good, however. Thomas, et al., found almost perfect interrater reliability for the APACHE II[5] and MEDISGRPS[6] systems, and relatively good reliability for the Clinical Staging system of SysteMetrics (614).

Type and source of hospital admission are sometimes used as proxies for patients' risk of dying (80,81,648). Coding of such information can be another source of error. The study by the California utilization and quality control peer review organization (PRO) of premature discharge notes that guidelines for admission source are subject to interpretation by coders (117). For example, it is unclear whether the referring physician or the transferring facility takes precedence. With transfer from another hospital a surrogate for patients' risk of dying (648), errors in coding source of admission could have affected hospital results. Some hospitals responding to the 1986 HCFA analysis commented that sources of admission had been recorded incorrectly by HCFA (648). Similarly, Blumberg eliminated 10 hospitals from his analysis of Maryland hospital data because they differed from other hospitals in the way they coded whether admissions were emergent, urgent, or elective (80,81). Only nonelective surgeries were included in Blumberg's study. If elective surgeries, which presumably entail less risk of death, were included for some hospitals and not others, the results would not have been valid.

Blumberg has noted a discrepancy between in-hospital deaths reported to State agencies and those reported to HCFA in 1984, with the number reported to HCFA lower than that reported to States (78). Similarly, a study by the California PRO found that 23 percent of cases that had been coded as being discharged alive from California hospitals had actually been discharged dead (117). The reasons for these errors are for the most part unclear; the California PRO study did find, however, substantial miscoding in the DRG series for patients with acute myocardial infarction. In that DRG series, Medicare payment for patients who are discharged dead is lower than payment for patients discharged alive.

Differing hospital policies concerning the point at which individuals are declared dead (141) and varying do-not-resuscitate policies do not affect the coding of death, but affect the reliability of patient death information across hospitals, which in turn affects the reliability of hospital mortality rates as an indicator of quality.

Statistical analyses should be validated with reviews of medical records. A significant problem in reviews of medical records has been interrater reliability (see ch. 7). The one published study of hospital mortality rates that addressed reliability among reviewers found good interrater reliability when reviewers used explicit criteria, but poor interrater reliability for subjective judgments of care (190). Other studies comparing explicit with implicit review have found similar results (see ch. 7).

---

[5] Acute Physiology and Chronic Health Evaluation.

[6] Medical Illness Severity Grouping System.

VALIDITY OF THE INDICATOR

Intelligibility of Hospital Mortality Rates as an Indicator of Quality

To be useful as an indicator of the quality of care, hospital mortality should be understandable to both consumers and providers. Anecdotal evidence indicates that consumers seem well aware that a patient's inherent risk of dying is a prime contributor to whether a patient lives or dies during or soon after a hospital stay. They also seem aware, however, of the hospital errors that can result in patient death.

For providers, examination of individual patient deaths may have face validity, but aggregate hospital mortality rates may not. According to Friedman and Shortell, mortality is the outcome that always receives the most intensive scrutiny by hospital managers and clinical chiefs of staff (237). Particularly in teaching hospitals, the medical staff discusses the causes of unexpected individual patient deaths (at least those deaths among patients of interns and residents) and suggests improvements in care. There is little evidence that hospital staffs examine overall hospital mortality rates or rates within hospital departments on a systematic basis (224). To date, providers have regarded skeptically attempts such as HCFA's to adjust statistically for patient characteristics that would explain high mortality rates so that the remaining explanation for differences among hospitals is the quality of care (97,537). It is unclear, for example, whether practicing physicians believe that a patient's likelihood of death can be predicted using systematic means. The use of mortality rates may be gaining in acceptance, however. The Commission on Professional and Hospital Activities (CPHA) reports in its mortality analysis that hospitals informally confirmed that high outliers had quality problems (172).

When To Measure Death

Researchers and policymakers do not yet (and may never) agree on when to measure an outcome of hospital care. Regional variations in lengths of stay among hospitals, differences in admitting and discharge practices, and unequal access to home care and hospice services in communities can determine whether a death occurs in the hospital or out of it (141). There seems to be considerable agreement that merely counting deaths at discharge is not a completely valid way to compare hospital mortality rates, because such a technique may reward hospitals that discharge patients in more serious condition, who may then die elsewhere. To capture a high percentage of deaths that may be attributable to poor-quality care, some analyses have used all deaths occurring within some time frame after an admission or a procedure, even if they did not occur in the hospital (588,589,647). This approach may, however, measure the effect of events unrelated to the quality of a hospital's care.

In empirical work relating to these issues, the Stanford Institutional Differences Study obtained essentially the same results from its Extensive Study (deaths at discharge only) as it did from its Intensive Study, which measured deaths at 40 days (even after discharge) or severe morbidity within 7 days of surgery (215,588). DesHarnais and her colleagues analyzed HCFA's 1986 data and found an almost perfect correlation between inhospital mortality rates and 30-day-post-admission mortality rates. DesHarnais and her colleagues concluded that "it does not matter which measure is used in terms of assessing hospitals' relative rankings" (172). It may be, however, that the conclusion would differ if all admissions rather than last admissions were included in the analysis. HCFA's 1986 analysis used patients' last admission of the year as the denominator in its analysis. A further consideration is that the appropriate time at which to measure outcome may vary for different conditions (500); this issue has not been tested.

For practical reasons, or because no valid endpoint has been established, most analyses have measured inhospital death only (80,81,189,190, 259,353,448,526,640). Clearly, this question requires careful thought and additional study.

Adjusting for Patients' Risk of Dying

One of the most challenging questions in quality assessment is how to construct an indicator that is not confounded with the characteristics of the patients who come to the hospital. In most analyses, the patient attributes used to adjust for the risk of dying have been only rough proxies for characteristics that may be better measured by physiologic values (see table 4-3), although the physiologic values that predict death are as yet unknown (596). Studies that use claims data alone are limited to the data elements present on claims, such as Medicare's UB-82. These claims indicate patient characteristics, such as age, sex, and race; the principal diagnosis for which the patient was admitted to the hospital and up to five secondary diagnoses; the principal procedure and up to three secondary procedures; some potential sources of admission; type of admission (emergency, urgent, elective, newborn); discharge sta-

Photo credit: Foster Daily Democrat

Age is at best a crude indicator of patients' inherent risk of dying.

tus (including dead or alive and, if alive, place discharged to); and other types of information less relevant to hospital mortality analyses (657).

Age may be the most frequently used adjustment for patient mix, and there is, of course, a correlation between age and the likelihood of death. However, the relationship is not completely linear (667), and age remains at best a crude indicator of a patient's health status or physiologic reserve (76). HCFA found, for example, that "average age of Medicare patients" was not a significant predictor of mortality at the hospital level of aggregation (640). The 1986 HCFA analysis used age groupings instead of average age of patients in the hospital. At the hospital level of aggregation, several age categories were statistically significant. In other studies using data at the patient level of analysis and more refined methods of adjustment, age has been found to be significant (353,588,589). Even if measured adequately, however, a number of studies have shown that age can also be a risk factor for inadequate or poor treatment (134,318,549,700a). Similarly, adjustments for sex, race, and socioeconomic status can mask an interaction between a patient characteristic and the provision of poor-quality care (191). Average length of stay (526, 640), for example, seems particularly invalid as a hospital level adjustment for patient risk. Longer lengths of stay can themselves be indicative of poor quality. The use of easily available discharge data to adjust for case mix is a threat to the validity of the hospital mortality measure, because a patient's risk of dying cannot be adequately inferred from diagnostic categories such as DRGs or ICD-9 codes (629,630).

Measures that rely at least in part on clinical data on admission would appear to have more validity than proxy measures such as age, sex, race, source of admission, and comorbidities (352,353). A recent review of the status of severity measures concluded that "although intrinsic biological severity may one day be measurable, currently it is an abstraction" (596), but some classification systems have reported good results (93,94,190,352,353). Williams was able to explain about 80 percent of the variance in neonatal mortality using a combination of birthweight, sex, race, and whether the birth was multiple (i.e., twins); by far the best explanatory factor was birthweight (545,702). Brewster's MEDISGRPS technique relies entirely on clinical findings, while Knaus' APACHE II method includes age and some comorbidities,[7] as well as clinical findings (see figure 4-1). Perhaps in line with the conclusion that measuring intrinsic biological severity is difficult, Brewster's results are not as impressive as Knaus'. Brewster's mortality results have been published for shortness of breath (93), abdominal pain, and chest pain (94) as reasons for admission; Knaus' for patients in the intensive care unit (352,353).

Even these patient classification systems may not be able to cope with the fact that the patient's condition may change during hospitalization regardless of the medical care provided. Having some clinical information about the patient's status on admission seems clearly better than relying on comorbidities and complications recorded after discharge, because existing coding schemes cannot clearly distinguish between comorbidities existing on admission and complications acquired as a result of hospital care. But a patient's status on admission to the hospital will not reflect changes in the patient's status that occur solely as a result of the trajectory of illness.

Appropriate measures of patients' risk of dying may differ considerably by disease category or patient condition. Measures that mix deaths of patients due to chronic or late-stage conditions with those of patients having more acute, less severe illnesses, and use one type of adjustment may not be nearly so valid as measures using either one or the other type of condition. Conclusions about the most appropriate aggregations and adjustments for patients' risk of dying are difficult to draw from existing studies because of the wide variation in methods used. Only one study has actually analyzed data for the hospital as a whole, with no conditions or patients excepted (189). Others have removed from consideration obstetric patients (259,526), or considered only elderly and disabled patients (640,647). The HCFA patient data base is composed primarily of patients aged 65 and over.

In general, however, analyses at the hospital-level of aggregation have been able to explain

Figure 4-1.—Scoring of Patients Under the APACHE II System for Classifying Severity of Disease

APACHE II score = Sum of A + B + C

A

Acute physiology score

The acute physiology score is the sum of points for 12 physiologic variables:

- temperature,
- mean arterial pressure,
- heart rate,
- respiratory rate,
- oxygenation,
- arterial pH,
- serum sodium,
- serum potassium,
- serum creatinine,
- hematocrit,
- white blood count,
- Glasgow coma score.

Each variable is scored from -4 to +4 points.

+

B

Age points

Age points are assigned to patients according to their age as follows:

≤44 - 0 pts
45-54 - 2 pts
55-64 - 3 pts
65-74 - 5 pts
≥75 - 6 pts

+

C

Chronic health points

For patients who have a history of severe organ system insufficiency or are immunocompromised, points are assigned as follows:

a. for nonoperative or emergency postoperative patients - 5 pts
b. for elective postoperative patients - 2 pts

SOURCE: Office of Technology Assessment, 1988; adapted from W.A. Knaus, E.Q. Draper, D.P. Wagner, et al., "An Evaluation of Outcome From Intensive Care in Major Medical Centers," *Annals of Internal Medicine* 104:410-418, 1986.

---

[7] Unlike the comorbidity measure used in most adjustment methods based on claims data, the comorbidities in Knaus' APACHE II scheme must have been present within 24 hours of hospital admission (352, 353).

more of the variation in mortality than have analyses at more condition-specific levels, although there have been rather high proportions of variation explained for certain conditions. Hospital-level analyses have accounted for between 35 and 93 percent of the variance. This is not surprising because random variation is less likely at the level of the institution.

Some differences in the amount of variance accounted for among diagnostic categories may be explained by the use of inappropriate variables to adjust for patients' risk of dying. Another potential explanation for differences among diagnostic categories is the extent to which medical care and its quality influence death rates. Therapy is unlikely to prevent the deaths of late-stage cancer patients, so not much variation is introduced by factors not accounted for in a regression equation (357). For early heart disease, on the other hand, good treatment does exist and its application does make a difference, so the patient's condition may account for little of the variation in patient mortality.

Validation of Hospital Mortality Rates Against the Process of Medical Care

The best way to establish hospital mortality rates as a valid indicator would be to demonstrate a link between the process of care and the outcome of death. Some studies have attempted to do this, with conflicting results. In response to HCFA's analysis of 1984 data, which showed 29 New York State hospitals as having higher than expected mortality rates, the New York State Department of Health conducted a regression analysis with its own set of adjustments for patients' risk of dying, modified from HCFA's 1984 model (462). New York State found fewer outliers[8] than did HCFA. The Department then had PRO personnel examine the medical records of patients in DRGs with mortality rates above the statewide average. The reviewers concluded that only about 3 percent of these cases had quality-of-care problems (278,461).

In 1987, New York State did not do a regression analysis, but compared the results of its targeting certain deaths for review to HCFA's analysis of 1986 data (279,461). In this comparison, only 1 hospital of the 10 identified by HCFA as being high-mortality outliers had quality problems using New York State's standards. In general, high outliers had fewer problems than non-outliers (279).

Dubois and his colleagues used both explicit and implicit review to determine whether quality problems existed in hospitals initially identified as high or low outliers using claims data (190). The explicit review compared the medical care provided (as reflected in the medical records) to a provisional list of criteria for quality of care in the management of patients. In the implicit review, experts read a summary of the patient's care and judged whether the death was preventable. Dubois's validation study is impressive because it was careful to test the possibility that factors other than quality, such as patients' characteristics related to their risk of dying, accounted for differences in mortality rates, for three of the most common causes of death (190).

Explicit review resulted in no apparent differences in numbers of preventable deaths, and implicit review found significant differences between high- and low-outlier hospitals in preventable deaths only for pneumonia, not for acute myocardial infarction or cerebrovascular accidents. After adjustment for differences in patients' risk of dying and for the fact that deaths were oversampled, however, the researchers found significant differences in preventable deaths between high and low outliers for cerebrovascular accidents and pneumonia. They estimated that 5 percent of patients with those conditions entering one of the high outlier hospitals would have a preventable death, compared to a 1-percent chance of preventable death in a low-outlier hospital. The authors concluded that adjustments using claims show some promise of identifying hospitals with variations in quality, although their study should be regarded as preliminary.

Knaus, et al., found that the best ranked intensive care unit in their study of 13 hospitals used significantly more therapeutic interventions than all the other hospitals (353). Knaus realized that the amount of treatment is not a good indicator of differences in quality; he examined the components of increased treatment at the "best" hospital, and found differences in the type of treatment provided.

---

[8] Outliers are hospitals that have mortality rates that are significantly either higher or lower than expected.

Somewhat similarly, the Stanford Institutional Differences Study included some crude indicators of the process of care (588,589,590). The process measures, all at the hospital level, were the rate of pathology reports, the rate of pathology reports showing the presence of disease, the rate of pathology reports showing no disease, and the autopsy ratio. The study found no significant relationships between inhospital death and any of the process measures. The study's original plan was to conduct a better validation study, but this plan was not supported because it was judged to be too lengthy and expensive (588).

The results of these studies should be regarded cautiously, however. Both New York State and Dubois and his colleagues used implicit review of records, which may be unreliable (190) (see ch. 7 of this report). New York State's targeted mortality study concentrated largely on surgical patients, while HCFA's analysis covered all reasons for admission. New York State's 1984 model adjusted for some factors that could have been related to quality of care.

Comparisons of Hospital Mortality Rates With Other Potential Measures of Quality

Some reviewers of the literature on hospital mortality have concluded that hospital mortality may have some validity as a quality indicator because mortality showed theoretically expected relationships with other potential measures of the quality of care. In a review of 18 studies of hospital mortality, for example, Fink, Brook, and Yano found that the following hospital characteristics were associated with better outcomes: frequency of performing a procedure, size, communication among staff, commitment of staff, clinical experience, board certification of staff, and teaching status (209).

Some of the studies reviewed for this report also examined relationships between hospital mortality and potential measures of quality other than hospital mortality, primarily structural measures. One study compared mortality to scales combining mortality and morbidity (588). The results of these analyses, shown in table 4-4, indicate some significant relationships between primarily structural measures of quality, defined quite variably among studies.

Comparison of Different Hospital Mortality Analyses

Hospital-specific mortality rates have been released by a variety of sources (55,77,81,115,116, 640,647). The New York State Department of Health also conducted two analyses in response to HCFA's releases; hospital-specific information related to these analyses were not released to the public (279,462). Some releases are of unadjusted mortality rates (55,115,116) and other analyses attempted to adjust for patient characteristics (80, 81,462,640,647). Comparisons of these analyses are instructive in several ways: they illustrate the different results obtained when mortality rates are analyzed specific to diagnoses or procedures versus aggregated by hospital; they show the potential importance of adjusting hospital mortality rates for patient characteristics; and they show the variation in results obtained when different risk-adjustment procedures are used.[9]

California

Several available data sets contained information on California hospitals: HCFA's releases of 1984 and 1986 adjusted data (640,647) and analyses by three newspapers of unadjusted data released by California Medical Review, Inc., the California PRO for fiscal year 1985 and 1985-86 (12,359,597). Because these sources differ in several ways, some variation in results is expected. In particular, the California PRO releases were completely unadjusted for patients' risk of dying. On the other hand, all releases pertained only to Medicare patients, and the years analyzed were contiguous, so one might expect some overlap in

---

[9] The analysis is also limited. It recorded only the presence or absence of a hospital on a particular list. Alternative approaches would rank the hospitals or use actual mortality rates or ratios, standardized in some way. However, the large number of comparisons might also preclude tests of statistical significance.

Table 4-4.—Comparison of Hospital Mortality With Structural and Other Outcome Indicators in Hospital Mortality Analyses Reviewed by OTA[a]

Study[b] Structural variable(s)	Variable significantly related to hospital mortality		Variable not related to hospital mortality
	Lower mortality	Higher mortality	
Roemer, et al., 1968 (526)			
1. Technological Adequacy Score[c]	+		
2. Hospital control:			
a. Voluntary			+
b. Proprietary			+
Roemer and Friedman, 1971 (525)[e]			
1. Medical staff organization:[d]			
a. Permissive control		+	
b. Medium control	+		
c. Strict control			+
Goss and Reed, 1974 (259)[f]			
1. Technological Adequacy Score[g]			
2. Hospital control:			
a. Municipal		+	
b. Voluntary			+
c. Proprietary			+
3. Teaching status:			+
a. Some commitment to teaching		+	
b. No teaching approval			+
c. Greatest commitment to teaching	+		
d. Hospital control and teaching status combined[h]			
Stanford Center for Health Care Research, 1974 (588), 1976 (589); Flood and Scott, 1987 (215)[l]			
Hospital Characteristics:[i]			
1. Medical staff structure (ES):[j k]			
a. Hospital-employed physician ratio		+	
b. Surgical-staff-to-patient ratio	+		
2. Nursing staff structure (ES):			
a. Proportion of part-time nurses	+		
b. Proportion of full-time nurses who are registered nurses	+		
c. Nurse-to-patient ratio	+		
3. Medical staff structure and nursing staff structure combined[m]			
Impact of Surgeons and Surgical Staff Organization (IS):[n o]			
1. Proportion of contract physicians	+		
2. Number of surgical specialties in the department	+		
3. Average percentage of practice conducted at the study hospital			+
4. Proportion of board-certified surgeons			+
5. Strictness of admission requirements for new members			+
Surgeon Characteristics:[n p]			
1. Percent of practice conducted at study hospital	+		
2. Number of residencies surgeon has completed	+		
3. Number of years in practice			+
4. Board certification			+
5. Surgical specialization			+
Within-Domain and Encroaching Influence (IS):[o q]			
1. "Control variables":[r]			
a. Percentage of surgeon's practice conducted at hospital			+
b. Hospital expenditures	+		
c. Patient's income	+		
2. "Power variables":			
a. Influence of the hospital administration within its own domain	+		
b. Encroachment by physicians on the nursing administration			+
c. Influence of the nursing administration within its own domain		+	

Table 4-4.—Comparison of Hospital Mortality With Structural and Other Outcome Indicators in Hospital Mortality Analyses Reviewed by OTA—Continued

Study[b] Structural variable(s)	Variable significantly related to hospital mortality		
	Lower mortality	Higher mortality	Variable not related to hospital mortality
d. Encroachment by physicians on the hospital administration's domain			+
e. Influence of the surgical administration within its own domain			+
f. Encroachment by the hospital administration on the surgical administration			+
Power of Surgical Staff Over Its Own Members (IS):[o,q]			
1. "Control variables":			
a. Percentage of surgeon's practice conducted at hospital			+
b. Hospital expenditures	+		
c. Patient's income	+		
2. "Power variables":			
a. Power of surgical staff over tenured surgeons	+		
b. Admission requirements for new members of the surgical staff	+		
c. Centralization of decisionmaking within the surgical staff			+
Selected Control Variables (IS and SIS):[s,t]			
1. Frequency of case discussions with pathologists			+
2. Control exercised by surgical staff over tenured surgeons			+
3. Chief of surgery's administrative influence in own area			+
NAS, 1977 (448)[u]			
1. General, urologic, and orthopedic surgeries combined:			
a. Degree of affiliation with a medical school			+
b. Proportion of surgeons who are board certified			+
c. Proportion of surgeons who are residents			+
d. Average age of surgeons			+
e. Absolute number of surgical beds			+
f. Proportion of acute-care beds allocated to surgery			+
g. Complication rate[v]			+
2. Cardiac surgery:			
a. Volume	+		
Knaus, et al., 1986 (353)			
1. Administration of unit, scope of service[w]			+[x]
2. Teaching status			+
Blumberg, 1987, 1988 (80,81)			
1. Teaching status			+

Study	Correlations among outcome variables (IS)	Outcome A	Outcome B	Outcome C	Intermediate scaled outcome
Stanford Center for Health Care Research, 1974 (588), 1976 (589); Flood and Scott, 1987 (215)	1. Outcome A: Death within 40 days of surgery (including after discharge)	—	Moderate positive correlation	Negative correlation	Small positive correlation
	2. Outcome B: Death within 40 days of surgery or severe morbidity at 7 days after surgery	—	—	Small positive correlation	Strong positive correlation
	3. Outcome C: Death within 40 days or surgery or severe or moderate morbidity at 7 days after surgery	—	—	—	Strong positive correlation
	4. Intermediate Scaled Outcome (ISC): Dead (9 points); severe (5) or moderate morbidity (2); else (0).	—	—	—	—

Table 4-4.—Comparison of Hospital Mortality With Structural and Other Outcome Indicators in Hospital Mortality Analyses Reviewed by OTA—Continued

[a] Only structural and outcome indicators are included in this table. Inclusion of process variables in studies is shown in table 4-3. Most of the analyses were a part of the primary publication reviewed by OTA. This table also includes, however, closely related studies using the hospital mortality indicator developed in the 13 analyses reviewed by OTA. For example, Roemer and Friedman (525) used the hospital mortality indicator developed in Roemer, et al (526).

[b] Numbers in parentheses refer to numbered entries in the reference list at the end of this report.

[c] The components of the Technological Adequacy Score used by Roemer and his colleagues were as follows, with points assigned to each component in parentheses:

1. Accreditation by the Joint Commission on Accreditation of Hospitals (20)
2. Approved residency or internship (10)
3. Approved cancer program (8)
4. Intensive care unit (7)
5. Pathology laboratory (5)
6. Blood bank (5)
7. Therapeutic X-ray (5)
8. Postoperative recovery room (5)
9. Rehabilitation service (5)
10. Outpatient department (8)
11. Home care program (8)
12. Social service department (7)
13. Chest X-ray on admission (7)

A total of 100 points could be scored. The source of data was hospitals' reports to the American Hospital Association.

[d] Roemer and Friedman devised a typology in which they defined medical staff organizations along a continuum from loosely structured or permissive to highly structured or vigorous (see Roemer and Friedman, 1971, ch. 5). Many of the components of the medical staff organizations were subsequently disaggregated in studies using data from the Stanford Institutional Differences Study (see Flood and Scott, *Hospital Structure and Performance*, 1987).

[e] Roemer and Friedman analyzed data for only 10 general hospitals in California but included a Veterans Administration hospital. Hospitals were chosen to represent a range of medical staff organization types, from loosely to most highly structured. Hospitals were also chosen to be generally meritorious.

[f] Goss and Reed used the same severity adjustment method as Roemer and his colleagues used.

[g] Goss and Reed used the same scale for the Technology Adequacy Score as Roemer and his colleagues used, except that "chest X-ray on admission" was omitted because data were not available.

[h] Municipal hospitals with internship and/or residency approvals had the highest severity-adjusted death rates; voluntary hospitals with medical school affiliations had the lowest death rates. No statistical tests were performed for any of the structural analyses.

[i] *Source*: A.B. Flood, W.R. Scott, and W. Ewy, "Hospital Characteristics and Hospital Performance," *Hospital Structure and Performance*, A.B. Flood and W.R. Scott (eds.) (Baltimore, MD: John Hopkins University Press, 1987).

[j] The initials ES, IS, and SIS in the entries that follow indicate whether the analysis was conducted with data from the Extensive Study (ES), the Intensive Study (IS), or the Service Intensity Study (SIS). The outcome in the ES and the SIS was in hospital death. The outcomes in the IS were, for the logistic regression, death within 40 days of surgery (including death after discharge), death within 40 days of surgery or severe morbidity at 7 days, death within 40 days or severe or moderate morbidity at 7 days. For the linear regression, moderate and severe morbidity at 7 days and mortality within 40 days were combined into a scaled measure. Only the Intermediate Scaled Outcome was used for most analyses (death [9 points], severe morbidity [5], moderate morbidity [2], and no or mild morbidity [0]).

[k] Adjusted for hospital size, teaching status, and expenditures, as well as patient characteristics.

[l] Note that in many of the analyses, hospital characteristics (size, teaching status, and expenditures) were controlled in addition for patient health characteristics.

[m] Results for medical staff and nursing staff combined were almost identical to those for individual variables, but when both sets were combined, the results for proportion of full-time nurses who were registered nurses were not significant.

[n] *Source*: A.B. Flood, W.R. Scott, W. Ewy, et al., "Effectiveness in Professional Organizations," *Hospital Structure and Performance*, A.B. Flood and W.R. Scott (eds.) (Baltimore, MD: John Hopkins University Press, 1987).

[o] Using the Intermediate Scaled Outcome (death [9 points], severe morbidity [5], moderate morbidity [2], and no or mild morbidity [0]).

[p] Aspects of hospital context (size, teaching status, and expenditures) were included in the analysis.

[q] *Source*: A.B. Flood and W.R. Scott, "Professional Power and Professional Effectiveness: The Power of Surgical Staff and the Quality of Surgical Care in Hospitals," *Hospital Structure and Performance*, A.B. Flood and W.R. Scott (eds.) (Baltimore, MD: John Hopkins University Press, 1987).

[r] Control variables entered into the analysis first.

[s] The outcome in the Service Intensity Study (SIS) was inhospital death.

[t] W.R. Scott, A.B. Flood, and W. Ewy, "Organizational Determinants of Services, Quality, and the Cost of Care in Hospitals," *Hospital Structure and Performance*, A.B. Flood and W.R. Scott (eds.) (Baltimore, MD: John Hopkins University Press, 1987).

[u] This study used basically the same method as the Stanford Institutional Differences Study (215,588,589), apparently without the admissions data.

[v] Data were collected for only 12 hospitals. The data were not routinely available in existing reports and the researchers were required to ask various hospital personnel for parts of the record. Further contributing to the possible lack of validity of this measure, the authors note that the definition of complication was somewhat subjective.

[w] Knaus, et al. based their designations of ICU levels on guidelines of the National Institutes of Health (NIH) Consensus Development Conference on Critical Care (334). The NIH Conference included variations in technological capability in its designation of levels. The hospitals in Knaus, et al.'s sample all had the same technological capability, however, so the assignment of levels was based on administrative structure only (353).

[x] The hospital with the lowest adjusted mortality rate was a Level I unit, and the hospital with the highest adjusted mortality rate was a Level III unit. As a group, however, Level I units did not do better than Level II or III units.

SOURCE: Office of Technology Assessment, 1988.

results. The appropriate comparisons are between HCFA's results for 1984 and 1986 and all other results, because data from the California PRO are actually for three different geographic areas.

With all sources and types of diagnoses and procedures combined, 143 (29 percent) of the approximately 490 California Medicare hospitals were either high- or low-mortality outliers in at least one analysis. Twenty-seven hospitals (19 percent of the 143 or 5 percent of all California hospitals) appeared as outliers in more than one analysis.

New York

As described above, New York State undertook two types of analyses to validate HCFA's releases. One was a regression analysis to detect outliers and the other was a targeted mortality analysis validated by PRO staff. The regression analysis was applied to the 1984 data, and the targeted mortality technique was applied to the 1986 data.

For the 1984 analysis, the New York State Department of Health used its own extensive data base to create predictor variables somewhat different from HCFA's, although like HCFA's, the analysis was conducted at the hospital level (462). In addition to identifying outliers for Medicare patients, New York State identified high-mortality outliers in 1984 for all patients under age 65, all discharges, and obstetrics/nursery services. The results of OTA's comparison indicate that 52 (19 percent) of New York State's 274 Medicare hospitals were high-mortality outliers on at least one of these 1984 analyses. Twenty-nine hospitals were HCFA outliers. But only 12 (23 percent) of the 52 high-mortality outliers were high outliers in both the 1984 HCFA aggregate analysis and at least one of the 1984 New York State analyses. Only half of these 12 were both HCFA and New York State Medicare outliers, indicating that New York State was able to replicate only 20 percent of the HCFA high-mortality outliers (6 of 29 HCFA high-mortality outliers) with its model. However, none of the 18 hospitals that were low-mortality outliers in HCFA's aggregate analysis were high-mortality outliers in the New York State analysis.

New York State used its targeted mortality method to critique HCFA's 1986 analysis (279). The targeted mortality study approach developed a set of case characteristics that are hypothesized to have a higher than average association with quality-of-care problems (461). New York State hypothesized that reviews targeted at cases rather than outlier hospitals would be more efficient at uncovering quality problems. In the New York State study, the targeting characteristics included procedures rarely associated with death, cases within DRGs that are rarely associated with death, cases in which the patient died in the hospital within 48 hours of surgery, surgical cases with a secondary diagnosis of acute renal failure, and cases with burns or poisoning as a secondary diagnosis. Cases meeting these screening criteria were forwarded for implicit review to a registered nurse; if the nurse concluded that the care provided either might not have met professional standards or might have contributed to the death or disability of a patient, the case was reviewed by one, and possibly two, physicians.

Comparing the results of the targeted mortality study with HCFA's 1986 analysis, New York State found only one high-outlier hospital in which there was a higher percentage of cases in which care either departed from standards or caused or contributed to patients' death than was found in the nonoutlier hospitals included in the study (279). In general, hospitals that were nonoutliers in HCFA's analysis had more quality-of-care problems than did outlier hospitals. The results of this study should be viewed somewhat cautiously, however, because the targeted mortality analysis used for comparison focused more on surgical than on medical cases, while the HCFA analysis covered all diagnoses.

It is striking that one hospital was a high-mortality outlier in almost all analyses. It did not show up as a problem in the 1984 obstetrics/nursery service data, however.[10]

Maryland

Maryland hospitals have received perhaps the most frequent examination of their mortality rates, although no attempt has been made at replication of specific analytic methods. The following analyses dealt with hospitals in Maryland: Bargmann and Grove (55), HCFA (640,647), Blumberg (80,81), and Washington Consumer Checkbook (693).

There is little convergence among the Maryland releases, which might be expected because of differences among analyses. Maryland has 58 Medicare hospitals (647). Of the 42 hospitals (72 percent) with actual mortality rates higher than those predicted by the various models, 17 hospitals (29 percent) appeared on more than one analysis. Seven of the 17 appeared as low-mortality outliers on one list and as high-mortality outliers on another list. Thus, only 10 (17 percent) appeared as high-mortality hospitals on more than one analysis, and their appearance was frequently for different procedure/condition categories. One hospital appeared as a Medicare outlier in both

---

[10] This hospital was also a high-mortality outlier in HCFA's 1984 analyses of DRG groups 089 (pneumonia), 115 (pacemaker implants), 121 (acute myocardial infarction), 127 (congestive heart failure), and 174 (gastrointestinal hemorrhage), but not DRG groups 106 (coronary artery bypass surgery), 195 (cholecystectomy), 209 (major joint procedures), or 336 (transurethral prostatectomy). Nor was it a low-mortality outlier in the latter groups. It was also a high-mortality outlier overall in HCFA's 1987 release of 1986 data.

1984 and 1986. It was also categorized as a high-mortality hospital for two procedures in Bargmann's and Grove's analysis (55).

Coronary Artery Bypass Graft Surgery in Arizona

Patten compared coronary artery bypass graft surgery (CABG) mortality rates in Arizona for two periods: when the Arizona certificate-of-need process was still in effect (July 1, 1984 to March 15, 1985), and after it was repealed (March 15, 1985 to December 31, 1986) (478). Table 4-5, which compares Patten's data with HCFA data, shows no overlap between HCFA 1984 CABG data (640) and that published for the two periods by Patten in the Phoenix *Gazette*. For the latter period, there was some convergence between the HCFA 1986 results (647) and Patten's results, although only one hospital was both a HCFA and a Patten outlier in 1986. In considering the 1986 data, one should keep in mind that in 1986 HCFA did not aggregate data by procedures, such as CABG.

Coronary Artery Bypass Graft Surgery in the District of Columbia

Three Washington, DC, hospitals had the three highest crude mortality rates for CABG in a Washington *Consumer Checkbook* analysis (out of 7 hospitals studied in the Washington, DC, metropolitan area), but no Washington, DC, hospital appeared on the list of 1984 HCFA outliers for CABG, or as an outlier for severe chronic and acute heart disease in 1986 (693).

Longitudinal Analyses

For hospital mortality as a measure of quality, it is important to know if a hospital's mortality rate in the past will predict its mortality in the future. However, there are many reasons why a hospital's mortality rate (and quality of care) may change over time, including random error in measurement; changes in the types of patients served; and changes in staff, practices, or procedures. Longitudinal studies are needed to gain insight into the likely role of random error in the

Table 4-5.—Comparison of Hospital Mortality Rates: Arizona[a]

	HCFA 1984		HCFA 1986			Health Services Advisory Group	
						Coronary artery bypass graft surgery	
	All diagnoses	Coronary artery bypass graft surgery	All diagnoses	Severe chronic heart disease	Severe acute heart disease	7/01/84 to 3/15/85[b]	3/15/85 to 12/31/86[c]
Hospital #1	—	—	—	—[d]	High[g]	N.A.[e f]	High[h]
Hospital #2	High[g]	—	—	—	—	—	—
Hospital #3	—	—	—	—[d]	—	N.A.[f]	High[h]
Hospital #4	High[g]	—	High[g]	—	High[g]	—	—
Hospital #6	—	—	—	—	—[d]	N.A.[f]	High[h]
Hospital #7	—	—	—	—[d]	—[d]	Low[i]	High[h]
Hospital #8	—	—	—	—	—	N.A.[f]	High[h]
Hospital #9	—	—	—	—	—	High[h]	High[h]
Hospital #10	High[g]	High[g]	—	—[d]	—	—	—
Hospital #11	—	—	—	—	—	High[h]	High[h]
Hospital #12	—	—	—	High[g]	—	—	—
Hospital #13	—	—	—	—	High[g]	—	—

[a] The HCFA mortality data referred to in this table were adjusted for patient characteristics such as age, sex, and comorbidities (see table 4-3). The Health Services Advisory data were not adjusted.

[b] Period when the certificate-of-need process was in effect.

[c] Period when the certificate-of-need process was no longer in effect.

[d] Near upper limit.

[e] N.A. = Not applicable.

[f] Did not perform open-heart surgery in this period.

[g] Mortality rate higher than expected.

[h] Mortality rate higher than State average.

[i] Mortality rate lower than State average.

SOURCES: **HCFA 1984 data:** U.S. Department of Health and Human Services, Health Care Financing Administration, "Medicare Hospital Mortality Information, 1984," Washington, DC: March 1986. **HCFA 1986 data:** U.S. Department of Health and Human Services, Health Care Financing Administration, *Medicare Hospital Mortality Information, 1986*, Washington, DC: U.S. Government Printing Office, Dec. 1987. **Health Services Advisory Group data:** B. Patten, "Open Market, Open Heart: Special Report," *The Phoenix [Arizona] Gazette*, p. A-1, Aug. 26, 1987.

results, but almost no quality assessment studies have compared hospital mortality rates over time. The Service Intensity portion of the Stanford Institutional Differences Study found little difference over time in service intensity or outcome (215,223), but this analysis was limited to 17 hospitals that volunteered to be in the study.

A summary of OTA's analyses of HCFA's data for 1984 and 1986 for the four States and the District of Columbia is shown in table 4-6. There was little convergence between the two HCFA results for these jurisdictions, but it is difficult to say whether these differences were due to actual changes in the hospitals, the fact that HCFA used different methods, or flaws in one or both of HCFA's methods. In California, for example, there were 18 HCFA high-outlier hospitals in 1984, and 20 in 1986, but only 2 (10 percent of the 1984 total) of the outlier hospitals were outliers in both 1984 and 1986. Another 37 hospitals, however, had actual mortality rates at or near the upper limit of the expected range of mortality rates in 1986; 7 of those had been outliers in the 1984 analysis. Thus, at most, 9 of the 18 1984 outliers in California(50 percent) had potential quality problems in 1986, if HCFA's methods are accepted as potentially valid indicators of quality problems. A number of other hospitals had actual mortality rates at or near the upper limit of the expected range of mortality rates in 1986, but were not high outliers in 1984 (647).

Table 4-6.—Number of Hospitals Found To Be High-Mortality Outliers by HCFA in 1984 and 1986, Selected States[a]

State	Number of hospitals		Convergence
	1984[b]	1986[c]	
California	20	18	2 of the 18 outliers in 1986 were also outliers in 1984
New York	29	10	5 of the 10 outliers in 1986 were also outliers in 1984
Maryland	1	1	The 1 outlier in 1986 was also an outlier in 1984
Arizona	3	1	The 1 outlier in 1986 was also an outlier in 1984
District of Columbia ...	0	1	The 1 outlier in 1986 was not an outlier in 1984

[a] High-mortality outliers are hospitals with mortality rates that exceed the expected range of hospital mortality rates. This table shows results for overall mortality rates only, not for specific diagnostic categories.

[b] U.S. Department of Health and Human Services, Health Care Financing Administration, "Medicare Hospital Mortality Information, 1984," Washington, DC: Mar. 10, 1986.

[c] U.S. Department of Health and Human Services, Health Care Financing Administration, *Medicare Hospital Mortality Information, 1986* (Washington, DC: U.S. Government Printing Office, Dec. 17, 1987).

SOURCE: Office of Technology Assessment, 1988.

New York had 29 high-outlier hospitals in HCFA's 1984 analysis of overall statistics and 10 in 1986; 5 (17 percent of 29) of the hospitals were outliers in both years. Another 21 hospitals were at or near the upper limit of the expected range in 1986. In Maryland, the same single hospital was a high-outlier hospital in 1984 and 1986. Eleven additional hospitals had mortality rates at or near the upper limit of the expected range of mortality rates in 1986. When all diagnostic categories are considered, Arizona had two outlier hospitals in 1984 and one in 1986. One (50 percent) of these hospitals appeared on both lists. None of the District of Columbia's hospitals were high outliers in 1984; one was in 1986.

A better longitudinal analysis of the HCFA data would be based on results using the same statistical method for the 1984 and 1986 data, as well as for 1985 data. HCFA conducted such an analysis using the 1986 analytical techniques and found a 44-percent convergence of high outliers between 1984 and 1986, and a 48-percent convergence between 1985 and 1986 (647). Neither the names of the hospitals that were outliers for any 2 of the years, nor any of the 1985 data were published, however, because the analysis did not generate hospital-specific information.

Although longitudinal data on hospital mortality may seem preferable to cross-sectional (onetime) data, consumers must be careful to consider changes in other factors that may occur over time that may affect the reliability of the indicator. For example, declining admissions as a result of policy changes can make mortality rates seem to change as well, because more severely ill patients may be admitted to the hospital (192,519). Consistent use of patient adjustments could alleviate this problem in the future, but in the past, different adjustments have been used in every release (640,647).

FEASIBILITY OF USING THE INDICATOR

Construction of the Indicator

Valid hospital mortality information depends on valid and reliable adjustments for the patient characteristics that increase the likelihood of death independently of the medical care provided. Conceptually, systems to adjust for patients' risk of dying based on clinical data on admission seem to be the most nearly valid means to adjust mortality rates. However, such systems are also the most costly to use and develop, primarily because they involve the collection of patient data not currently in the discharge abstracts routinely compiled by hospitals for billing purposes.

Adjustment of hospital mortality rates may also involve the calculation of separate algorithms for individual diagnoses and procedures; these algorithms will need to be continuously updated if they are to conform to advances in statistical methods and medical practice. Such efforts will require expertise in statistical and research methods and medical practice. Similar efforts are required to devise ways of comparing hospital mortality rates to the process of care (442).

The Joint Commission on the Accreditation of Healthcare Organizations' clinical indicators project is assessing the feasibility of regularly collecting clinical data. Preliminary estimates indicate that the collection of such data will be relatively expensive. It will also be expensive to check mortality data against process of care information. Both the Maryland Hospital Association and CPHA have provided member hospitals with workbooks containing hospital mortality norms.

A number of providers and consumer representatives have suggested that analyses within clinically meaningful diagnosis and procedure groups are both more nearly valid and more meaningful to individual consumers than hospital-wide mortality rates (500). However, results aggregated by hospital may be useful for evaluating institutional performance. Individual consumers may not be sophisticated enough to distinguish among hospital services. Finally, organizational purchasers of care may contract with entire hospitals, although they do sometimes contract for specific services from different hospitals. Thus, information at the hospital level would be useful to them. There clearly seems a place for information aggregated at different levels.

Intentional Manipulation of the Indicator

The fact of death does not seem easy to manipulate intentionally, but a focus on death rates without adequate research attention to adjustments for patients' risk of dying and validation against the use of appropriate medical processes may lead hospitals to refuse to accept severely ill patients, to postpone their admission from the emergency room, to discharge them hurriedly to other facilities, or to intentionally miscode diagnoses.

The California PRO study found substantial miscoding in a DRG series that had higher paid DRGs for patients discharged alive (117). On the other hand, neither the California PRO study nor

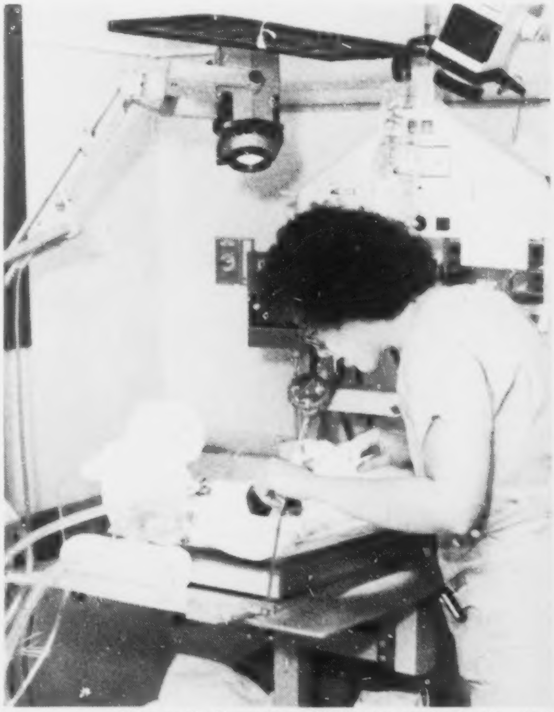


Photo credit: Strong Memorial Hospital, Rochester, New York

Consumers may wish to have hospital mortality information that is specific to particular conditions or services, such as neonatal intensive care.

an OIG study found a pervasive pattern of premature discharges within their definitions (117,660a). The OIG study covered only an early period of prospective payment implementation (October 1984 to March 1985), however, and also found that one in every five hospitals reviewed had at least one occurrence of a premature discharge; the occurrence was one in three in rural hospitals. The California PRO study, begun early in 1986, did find a higher proportion of premature discharges among patients who *died* within 20 days of discharge compared with premature discharges who were *readmitted* within the same period, and a significant pattern of premature discharges in patients readmitted within 1 day of discharge.[11] Thus, there is some evidence of premature discharge.

Some have suggested that earlier discharges of patients who are likely to die is one way to contend with the release of hospital mortality statistics, although they implied that the release be medically appropriate and to an appropriate alternative care facility (224,660a). There are also incentives and analytic procedures which may make premature discharges in the face of mortality releases unlikely, such as intensive PRO review and the use of 30 days post admission as the time when deaths are counted.

Another way for a hospital to reduce its mortality rate is to keep severely ill patients in the emergency room rather than admitting them to the hospital. Currently, emergency room patients are not counted as hospital admissions.

Other quality assessment/assurance mechanisms may have to be in place to prevent refusals to admit or premature discharge. For example, hospitals that participate in Medicare and transfer uninsured emergency patients before stabilizing their conditions will be fined $50,000 (Public Law 100-203). Another way to discourage hospitals from transferring patients or discharging them prematurely would be to "credit" each hospital that sees a patient during an episode of care for the patient's death.

Dissemination of Information About the Indicator

Hospital-specific mortality rates are becoming more available to the public (see box 4-A). HCFA's releases of 1984 and 1986 Medicare data were, of course, the most prominent. The California PRO released mortality information about California Medicare patients in 1986 and 1987. The University of California Santa Barbara has available for sale hospital-specific data on infant mortality from its Maternal and Child Health Data Base (598). Blumberg's analysis is available to the public. Portions of all of these reports were reported in newspaper articles and in reports by consumer groups (503,598,693). In addition to information made available to the public, the Maryland Hospital Association and CPHA calculated hospital mortality norms and made the information available to their member organizations (141,408).[12]

Consumer advocates have applauded the availability of hospital mortality data (115), but reservations have been expressed as well (14,500). Some hospitals are reported to be making mortality data available as part of a marketing strategy (426), while others continue to criticize the release of such data in its present state (41, 426,427). Some States have mandated the collection and reporting of numerous clinical indicators and are also planning to release outcome data (41,427). The cumulation of information in its current methodological state may be helpful to consumers who are relatively sophisticated. Those who are less sophisticated will probably need a considerable amount of help interpreting the data.

HCFA used the media to disseminate to the public the 1984 and 1986 hospital mortality rates for Medicare patients (648). HCFA was reluctant to publish the 1984 data, but was pressed to do so by the possibility of a Freedom of Information Act suit (302). In the press release accompanying the 1986 analysis, HCFA characterized the release

---

[11] The California PRO study is flawed in that it looked only at these two groups. Presumably, there may have been patients discharged prematurely who were neither readmitted nor died soon after discharge. The OIG study recognized this problem and identified premature discharge regardless of whether the patient did well subsequently.

[12] CPHA's workbook is available for sale to the public as well, although it is not clear how useful it would be to the general consumer.

Box 4-A.—Selected Sources of Information About Hospital-Specific Mortality Rates

Type of information	Source(s)
National information about mortality rates among 1984 Medicare patients	Director, Health Standards and Quality Bureau Health Care Financing Administration U.S. Department of Health and Human Services 6325 Security Boulevard Baltimore, MD 21207
National information about mortality rates among 1986 Medicare patients	1. *Medicare Hospital Mortality Information,* Stock No. 017-060-00206-9 U.S. Government Printing Office Washington, DC 20402 Cost: $69 for a 7-volume set 2. American Association of Retired Persons regional offices; Main office: 1909 K Street, N.W. Washington, D.C. 20048 3. Utilization and quality control peer review organization (PRO) offices
Information on mortality rates among California Medicare patients for 24 diagnostic categories, April 1, 1985 through March 31, 1986; and for the 50 most common diagnosis-related groups, Federal fiscal years 1985 and 1986	California Medical Review, Inc. 1388 Sutter Street Suite 1100 San Francisco, CA 94109 Telephone: (415) 923-2000 Cost: $10 for each hospital listing
Information on California perinatal mortality rates, 1980-84	Maternal and Child Health Data Base c/o Community and Organization Research Institute University of California, Santa Barbara 2201 North Hall Santa Barbara, CA 93106 Cost: Descriptive narrative, $20; statistical appendix, including individual hospital statistics, $30[1] [2]
Information on mortality rates in Maryland for nine surgical categories, 1979-80	*Surgery in Maryland Hospitals 1979 and 1980: Charges and Deaths,* by E. Bargmann and C. Grove, 1982 Public Citizen Health Research Group 2000 P St., N.W. Room 708 Washington, DC 20036 Telephone: (202) 872-0320
Information on mortality rates in Maryland for nonelective surgeries, April 1984-March 1985[3]	Maryland Health Services Cost Review Commission 201 W. Preston St., First Floor Baltimore, MD 21201

---

[1] Source: R. Steinbrook, "Hospital Death Rates: A Wide Variance," *Los Angeles Times*, p. 1, June 15, 1987.

[2] A summary of these data for hospitals in the Los Angeles area (including hospitals with high mortality rates in Riverside, San Bernardino, San Diego, and Ventura counties) was published in the *Los Angeles Times* in November 1987 (R. Steinbrook, "Care for Newborns Varies, Studies of Hospitals Show," *Los Angeles Times*, p. 1, Nov. 9, 1987).

[3] Mortality rates for digestive operations were reported in the *Baltimore Sun* in November 1987 (M. Knudson, "Death Rate Found To Vary for Digestive Operations," *Baltimore Sun*, p. 1A, Nov. 22, 1987).

SOURCE: Office of Technology Assessment, 1988.

as an "important contribution to the existing body of knowledge about health care" but stressed that a significant shortcoming of its approach was that it did not contain an objective and direct measure of the condition of the patient at the time of hospital admission. HCFA urged health care consumers to read the explanations of the information's uses and limitations, as well as comments provided by the hospitals, and when appropriate to discuss the indicators with their physicians or with hospital administrators. HCFA also cautioned that the information should not be used to rank hospitals, and was not designed to provide a national benchmark for measuring the quality of care. An additional set of questions and answers accompanying the press release cautioned the media not to report the mortality statistics as definitive measures of the quality of care (648).

For the 1987 release, OTA examined clippings from newspapers with circulations of 50,000 or more, and transcripts from radio broadcasts to review the manner in which the HCFA data were being reported. Burrelle's clipping service found about 100 clippings (including 3 radio transcripts and several editorials, op ed pieces, and letters to the editor). Fifty-two articles describing the HCFA release were written in newspapers in 19 States and the District of Columbia.[13] Typically, stories focused on hospitals that were outliers, and stated that area hospitals were either high, average, or low on the HCFA lists. The media quoted from the HCFA press release and/or HCFA personnel about the limitations of the study, and quoted local hospital or hospital association personnel, who generally criticized the release. There was very little understanding evinced that HCFA did try to adjust for patients' risk of dying using proxy measures, such as comorbidities, transfers from other hospitals, and previous hospitalizations. Of the 10 editorials, 6 were in favor of the data release, even if only because it made some information available to consumers, and 4 opposed the release. Stories about the release of the Dubois study a week after the HCFA release gave the HCFA release more credibility as a quality indicator (190).

Only one story gave consumers information about how to get the HCFA release itself. HCFA's press release had said that the report was available for $69 from the U.S. Government Printing Office (648). The report is sold as an entire set of seven telephone-book sized volumes with information for every State, the District of Columbia, American Samoa, and Guam. HCFA reports a steady, but not large, stream of requests for information about how to get access to the report (98). By May 15, 1988 (5 months after the report was released, and the last day for which data are available), the Government Printing Office had sold 236 sets (670).

Hospital mortality releases are costly. The University of California's report on perinatal mortality in California hospitals cost $50 in 1987 (598), and each individual hospital report from the California PRO cost $10 in 1987. HCFA sent copies of *Medicare Hospital Mortality Information, 1986* to State health offices, PROs, HHS regional offices, and 10 American Association of Retired Persons regional offices, and suggests to consumers that they also contact depositary libraries for copies of the report. HCFA acknowledges that not all of the sites they suggest for consumers to review the report are accessible to all people and not all of the sites will have copies of the report (98).

Both HCFA and the California PRO have suggested that concerns about high hospital mortality rates be the occasion for consumers to ask their physicians questions about specific hospitals (115,116,640,647). The Public Citizen Health Research Group has suggested some questions that consumers might ask, and advised that consumers be sure to get "specific and substantial answers" (503). Hospital comments on the HCFA release and other releases suggest that hospitals may respond by citing inaccuracies in the data, large numbers of admissions from the emergency room, patient characteristics, and patient and family wishes to not resuscitate, rather than errors in care, as explanatory factors (598,647).

---

[13] Twenty jurisdictions ran articles on the mortality release: Alabama, Arizona, California, Colorado, District of Columbia, Florida, Georgia, Louisiana, Maryland, Massachusetts, Missouri, New Jersey, New York, Ohio, Oregon, Pennsylvania, Rhode Island, Tennessee, Texas, and Virginia.

CONCLUSIONS AND POLICY IMPLICATIONS

Given the methodological and conceptual problems associated with using hospital mortality rates as an indicator of the quality of care, the rates cannot at present be considered definitive indicators of quality. The release of hospital-specific mortality rates does have the potential, however, to initiate a dialogue between consumers and providers. Such releases can also provide national and regional information against which hospitals and others can compare results; this information may lead to identification and correction of quality problems. Physicians, hospital staff, consumers, and organizations that contract with hospitals to provide care may be able to use hospital mortality rate information as leverage to improve care.

One of the advantages of using hospital mortality rates as an indicator of the quality of care is that it makes sense to average consumers. Consumers are also aware that patients' inherent risk of dying is usually the most important determinant of whether patients die during or soon after a hospital stay, and see the importance of adjusting for this risk.

One of the weaknesses of hospital mortality rates as a quality indicator is the current lack of methods to adjust adequately for patients' risk of dying in most analyses. In addition, methods for reviewing medical records to determine quality problems and validate statistical analyses are underdeveloped. The data on which analyses are performed may not be uniformly coded and collected, leading to spurious differences among hospitals. Hospital mortality rates are just beginning to be validated against the medical care provided. If mortality statistics are not valid but are nonetheless interpreted as indicators of the quality of care, their release could result in a breakdown of trust between patients and providers, and loss of reputation, patients, and income by hospitals.

Another problem may be that information on hospital mortality may not be available when consumers need it. An additional problem may be that individual consumers do not yet know either the appropriate questions to ask about hospital mortality statistics or the responses to accept. Information about hospital mortality rates for specific diagnoses may be as useful as information about overall institutional performance, but diagnosis-specific information is more difficult to obtain than information about the hospital as a whole. Most newspaper reports of the 1986 HCFA analysis did not include diagnosis-specific information. In any case, the diagnostic groupings included in the HCFA report may be neither understandable to consumers nor clinically meaningful in themselves.

Given the undeveloped state of hospital mortality statistics and the skepticism of some providers about the value of the information, it seems that consumers could use additional guidance about the kinds of questions to ask and the kinds of answers to accept when faced with anomalous mortality statistics. As a first step, it might be useful to develop information about the ways consumers and providers have responded so far to the mortality rate information that is already available. Are they aware of it? Are they frustrated by advice to regard the information cautiously or by hospitals' denial of its importance? The questions consumers ask about the mortality data and the responses the consumers were given could be studied. Perhaps more useful to consumers would be for HCFA to assess the validity of the mortality data so that quality concerns can be verified or dismissed. There was some consideration given to having PROs investigate the validity of the 1986 data (503), but ultimately they were not required to do so. Confidence in the results of any review would depend on the rigor of the review process.

Considerable progress seems to have been made in the development of valid hospital mortality indicators, but researchers, policymakers, and providers agree that considerable problems remain. HCFA's 1986 mortality rate analysis was much improved over its 1984 analysis, although neither has yet been validated against an independent criterion. Dubois's study is the most careful so far in its testing of alternative hypotheses to account for mortality outliers; his finding that outlier status based on claims data for the entire hospital overall was confirmed for two conditions suggests that the use of claims data to identify hospitals with quality problems has potential (190). The results of the study by Dubois, however, can be regarded only as preliminary, particularly because of the relatively low reliability among reviewers using implicit criteria, and the fact that only three conditions were included in the research. It is also important to keep in mind that the adjustment procedure used by Dubois and his colleagues was different from those used by HCFA and by other researchers, so that Dubois's results may not be generalizable to other methods.

There are a considerable number of methodological and conceptual issues to be resolved before hospital mortality rates can be regarded as a valid indicator of the quality of care. The resolution of these issues requires a comprehensive and quite large and costly research program. The construction of valid adjustments for patients' risk of dying and establishment of a link between mortality and the process of care will depend on the establishment of an orderly, iterative process (442). Most importantly, links to process must be established, and valid techniques found to adjust crude mortality rates for patient characteristics. Needed is additional attention to research design, statistical analysis (for regression studies), and methods for confirming quality-of-care problems. Ways to link patient information across different data sets are needed. For example, ambulatory files may provide useful information about patients' status on admission to a hospital. Given the results of the Dubois and New York State studies, ways to develop explicit criteria or increase the reliability of implicit review are critical. It may not be too early to develop conventions so that the more reliable and validated sources of data and methods are used. A number of studies related to these issues are in progress (see app. E). A study of nonintrusive outcome measures, being conducted by the Rand Corporation, is examining the relationship between prior medical care and death. HCFA is comparing the results of its 1986 analysis to one using MEDISGRPS to adjust for patients' severity of illness (357).

The validation process will be expensive. Certain types of experts are needed to develop adjustment methods, and many more experts are needed to review medical records and to establish the reliability of data bases. Even so, given the limitations in ability to measure patients' risk of dying, it may be that hospital mortality will continue for quite some time to be useful only as a screen or flag for possible quality problems.

In the absence of a validated method for constructing hospital mortality as a quality indicator, is the release of the information to the public justified? Brook and Lohr suggest that it is inappropriate to identify outlier hospitals publicly before evaluating the reliability and validity of the data and giving those hospitals adequate time to review their own data (104). Brook and Lohr suggest further that outlier hospitals be given up to 6 months to correct any problems before information is released. This approach would seem to encourage a closer working relationship between releasing bodies and the hospital, and perhaps more support for the release of data by the hospitals. If, however, as HCFA seems to intend (648), hospital mortality rates for all hospitals continue to be published, HCFA might follow up with reviews of the medical care process, so that the public would know whether quality problems were in fact confirmed. To do this, HCFA would need to develop a standard review method.

Releasing hospital mortality data may provide an incentive for hospitals to look more closely at the care they provide. A recent survey of hospitals showed little use of comparative death rates (132). Hospitals that conduct appropriate investigations of the reasons for differences may be able to improve the quality of care that they deliver. Physicians and organizations may find the data useful in referring patients or selectively contracting with hospitals. The rates of preventable deaths and the percentage of quality problems found in numerous studies (190) suggest that additional attention to patient care is warranted. Finally, once validated, adjusted hospital mortality rates and other outcome measures could be a good complement to studies based on reviews of medical care provided (process studies), and provide a good validation criterion for studies of structural properties of hospitals.

Chapter 5

Adverse Events

CONTENTS

Page
Introduction 101
State- and National-Level Occurrence Screens 106
Reliability of the Indicator 108
  Nosocomial (Hospital-Acquired) Infections 108
  HCFA's Generic Quality Screens 110
Validity of the Indicator 111
  Nosocomial (Hospital-Acquired) Infections 111
  Occurrence Screens 112
Feasibility of Using the Indicator 114
  Nosocomial (Hospital-Acquired) Infections 114
  Occurrence Screens and Incident Reporting 115
Conclusions and Policy Implications 116

Boxes

Box *Page*
5-A. Mandatory Incident Reporting in Massachusetts 104
5-B. Mandatory Incident Reporting in New York 105

Tables

Table *Page*
5-1. General Outcome Screening Criteria for Hospitals 102
5-2. JCAHO Hospitalwide Clinical Indicators Being Evaluated as Screens for Hospital Quality Problems 107
5-3. HCFA's Generic Quality Screens 108

Chapter 5

Adverse Events

INTRODUCTION

The idea that problem medical care can be identified through poor patient outcomes that are unexpected is behind the "occurrence screening" and "incident reporting" systems that have been implemented in almost all U.S. hospitals. Touted as early warning systems for hospital administrators, occurrence screening and incident reporting systems grew out of the malpractice crisis of the mid-1970s, when institutions desperately began to seek ways to limit their liability. Exactly what constitutes an occurrence or an incident varies widely among institutions. Although most reporting systems use patient outcomes as criteria to screen for occurrences or to define incidents, some also use criteria related to the process of care. The single thing that all the reporting systems have in common is that they are used by hospitals only as a first step for finding poor-quality medical care. In many cases, the occurrence of adverse events may result from factors other than poor quality. Thus, to establish a link between the quality of hospital care and adverse events, hospital cases identified by the reporting systems must be followed up with more thorough investigation and interpretation by medical advisers.

In the early 1970s, Rutstein and his colleagues proposed counting "sentinel health events," or cases of unnecessary diseases, disabilities, and untimely deaths, to monitor the quality of medical care (546). Working with numerous specialists, these researchers developed a list of specific conditions for which adverse outcomes—whether caused by commission or omission—should never occur, such as death from tuberculosis.

Specific criteria for reporting adverse incidents across all conditions were first developed in 1976 in the California Medical Insurance Feasibility Study (432). That study, sponsored by the California Medical Association and the California Hospital Association, used general outcome criteria to screen more than 20,000 patient charts from 23 hospitals for adverse events that might result in litigation for malpractice compensation. The 20 "potentially compensable events" developed by physicians and medical audit experts in the 1976 California study later became the basis for "occurrence screens," adapted and modified for use by individual institutions. An adaptation of the general outcome criteria that was developed by Medical Management Analysis is shown in table 5-1 (154). The outcome criteria in the table, now used in more than 200 U.S. hospitals, cover all aspects of hospitalization and are generally used to screen every patient record during the patients' hospital stay (290).

Among the common adverse events used as criteria in most occurrence screens are deaths, nosocomial (hospital-acquired) infections, unusually long lengths of stay, and unscheduled procedures, readmissions, or transfers. The use of deaths as a criterion may be limited to cases where death is a statistically rare outcome for the procedure, condition, or diagnosis-related group or is in some other way unexpected. In most hospitals, cases with adverse events identified by an occurrence screen are subsequently reviewed in depth for possible problems related to the quality of care. Almost all hospitals adapt occurrence screens for their own particular needs, for example, adding suitable clinical indicators developed at the departmental or service level. The use of occurrence screens in a hospital is usually part of the hospital's quality assurance program and therefore directly linked with existing peer review endeavors.

It is not known how many U.S. hospitals currently use occurrence criteria to screen their patient populations for adverse events. Increasingly, insurance companies are requiring hospitals to use occurrence screens as a condition for underwriting the medical malpractice insurance of affiliated physicians (420). The Joint Commission on the Accreditation of Healthcare Organizations (JCAHO) encourages the use of specific criteria to select cases for review in hospitals' quality assurance programs, yet it gives ample leeway in

Table 5-1.—General Outcome Screening Criteria for Hospitals

Criterion 1: Admission for adverse results of outpatient management.	**Exceptions:** Specific instructions may be developed by the clinical departments concerning expected admissions for chronic conditions managed in the outpatient setting.
Criterion 2: Readmission for complications or incomplete management of problems on previous hospitalization. a. Pre-existing complication with deterioration. b. New complication. c. Recurrent disease state. d. Unresolved disease state.	**Exceptions:** • Complication or incomplete management occurred at another hospital not associated with this hospital or involved a practitioner who is not on this medical staff. • Planned admissions for secondary procedures needed to complete treatment.
Criterion 3: Operative/invasive procedure consent. a. Incomplete. b. Missing prior to procedure. c. Different procedure done from procedure on permit. d. Different surgeon performed procedure than name on permit. e. Not signed by patient or legal guardian. f. No informed consent note. g. Other.	**Exceptions:** • Emergency procedures where the patient was unable and the family or legal guardian unavailable to sign the consent. • Life-threatening problems found and addressed during surgery.
Criterion 4: Unplanned removal, injury, or repair of organ structure during surgery or other invasive procedure, or vaginal delivery.	**Exceptions:** None.
Criterion 5: Unplanned return to operating room, delivery room, or other special procedures room on this admission.	**Exceptions:** Planned second procedure or second stage of a procedure planned prior to first procedure.
Criterion 6: Surgical and other invasive procedures which do not meet criteria for necessity and appropriateness. a. Diagnostic tissue—pathology report does not match preoperative diagnosis. b. Nondiagnostic or normal tissue removed and medical staff criteria for necessity or appropriateness not met. c. No tissue removed and medical staff criteria for necessity and appropriateness not met. d. Other.	**Exceptions:** As developed by the medical staff.
Criterion 7: Blood loss excessive or blood/blood component utilization which is unjustified, excessive, results in patient injury, or is otherwise at variance with professional staff criteria. a. Excessive blood loss occasioned by iatrogenic bleeding or anemia with or without transfusion. b. Transfusion of blood or blood components not clinically indicated. c. Transfusion reaction. d. Other.	**Exceptions:** As developed by the professional staff.
Criterion 8: Nosocomial infection (hospital-acquired infection).	**Exceptions:** Infection acquired outside this hospital, clinic, or home health care setting and did not involve any member of this medical staff.
Criterion 9: Drug/antibiotic utilization which is unjustified, excessive, inaccurate, results in patient injury, or is otherwise at variance with professional staff criterion. a. Does not meet professional staff criterion for appropriateness. b. Inadequate/excessive/inappropriate/inaccurate dosage or timing. c. Drug or contrast material reaction/interaction. d. Other.	**Exceptions:** As developed by the professional staff.
Criterion 10: Cardiac or respiratory arrest/low Apgar score.	**Exceptions:** None.
Criterion 11: Transfer from general care to special unit.	**Exceptions:** Transfer scheduled prior to surgery or other special procedure.
Criterion 12: Other patient complications.	**Exceptions:** None.
Criterion 13: Hospital-incurred patient incident. a. Falls, slips, patient accident. b. Intravenous problems, such as calculation errors, overloads, or infiltrations. c. Skin problems, such as rash, threatened or new decubitus ulcer. d. Equipment failures/malfunctions. e. Other incidents, such as procedural errors, electrical shock or burn, actual or attempted suicide, and lost or damaged property.	**Exceptions:** None.

Table 5-1.—General Outcome Screening Criteria for Hospitals—Continued

Criterion 14: Abnormal laboratory, X-ray, other test results, or physical findings not addressed by physician.	**Exceptions:** As developed by the professional staff.
Criterion 15: Development of neurological deficit which was not present on admission.	**Exceptions:** As developed by the medical staff for expected outcomes, such as deficits following intracranial surgery.
Criterion 16: Transfer to/from another acute care facility. a. Financial reasons. b. Management/procedures not available at this institution. c. Patient option. d. Other.	**Exceptions:** Mandatory transfer for administrative reasons, or transfer for tests not available at this hospital.
Criterion 17: Death a. Unexpected with surgery. b. Unexpected without surgery. c. Expected, disease related. d. Other.	**Exceptions:** None.
Criterion 18: Subsequent visit to emergency department or outpatient department for complications or adverse results from a previous encounter.	**Exceptions:** Planned returns for wound checks or suture removal.

SOURCE: J.W. Craddick, *Medical Management Analysis Series: Vol. II, Improving Quality and Resource Management Through Medical Management Analysis* (Rockville, MD: Medical Management Analysis International, Inc., 1987).

the degree of specificity. There are certainly wide disparities in the occurrence screens used by hospitals. All or samples of patient populations can be screened for occurrences either during the patients' hospital stay or retrospectively after discharge. Hospitalwide, "generic" screens can be applied equally across all patients, or detailed service-specific criteria devised for similar sets of patients. Screens can be computerized too, but the level of patient information detailed in the screens usually requires the review of patients' medical records by specially trained personnel in all but a few highly computerized hospitals.

Incident reporting systems, though often overlapping with occurrence screens and also growing out of concerns about rising malpractice liability, tend to be organized and operated somewhat differently from occurrence screens. Incident reporting systems are organized directly by the hospital administration (rather than being part of a hospital's quality assurance program) and tend to be operated independently of the medical record or other existing information systems. Typically, as part of risk-management programs, hospital personnel (most frequently nurses) complete forms when they observe an adverse event, and the forms are reviewed centrally by a hospital administrator/risk manager. The definition of an "incident" is often left to the discretion of the frontline health professionals who deal with patients. Most commonly, adverse events such as patient falls, medication errors, equipment failures, and commission of procedure or treatment errors are considered incidents. Reliance is placed on educating nurses, physicians, and other health care workers to recognize problems and report them.

Because health care personnel use their judgment in reporting incidents, it is more likely that incidents reflect quality-of-care problems than do the adverse events that are initially identified by occurrence screens; screening systems are expected to identify substantial numbers of false positives. Although reported incidents might therefore be viewed as being one step closer to identifying poor-quality care than are occurrences picked up in screens, further investigation of incidents is also necessary. First, an incident may not have been caused by negligent medical care; for example, a patient fall may have resulted from the patient's own carelessness. Second, an incident may not have had an important impact on the patient; for example, even though a medication has been administered incorrectly, a patient may suffer no ill effects.

Almost all hospitals have incident reporting systems, but the quality and reliability of reporting in these systems vary enormously across institutions. Currently, eight States[1] and the Veterans

---

[1] The eight States are Alaska, Florida, Kansas, Maryland, Massachusetts, New York, Rhode Island, and Washington (290).

Box 5-A.—Mandatory Incident Reporting in Massachusetts

Since July 1, 1987, all hospitals, clinics, and health maintenance organizations in Massachusetts have been required to submit detailed quality assessment plans—which must include reporting systems for both incidents and occurrences—to the Massachusetts Board of Registration in Medicine (the Medicine Board). State regulations, which grew out of the Malpractice Tort Reform Act of 1986, empower the Medicine Board (which also has responsibility for licensing and disciplining physicians) to approve or disapprove these quality assessment plans.

Health care institutions are required to submit copies of their occurrence screens and information on how the screens are to be used in their quality assurance programs to the Medicine Board, but they are not required to report the numbers or kinds of occurrences. (All the underwriters of physicians' malpractice insurance also require that hospitals use occurrence screens.) Likewise, all health care facilities must submit their plans for incident reporting systems to the Medicine Board. Summary reports of incidents must be reported to the Medicine Board at least quarterly.

Four major incidents have been defined in the Massachusetts regulations, and their reporting is mandatory for all providers: 1) maternal deaths related to delivery; 2) fetal deaths (excluding abortions); 3) chronic vegetative state resulting from medical intervention (the Medicine Board is refining this definition further at the complaint of the medical profession); and 4) death in the course of or resulting from ambulatory surgical care. Major impairments or deaths that are unexpected are also supposed to be reported, although their definition is left to the providers (243 CMR 3.08 (1987)). Reports on these incidents must include identification of the provider, a brief description of the incident, and patient data. Health care organizations also must define further criteria for incidents, but the ongoing reporting of other incidents is required only in summary form.

Because the system is so new, the Medicine Board has not as yet started to audit hospitals and other providers based on the incident reports (420). Although the right of the Medicine Board to collect and act upon the information was upheld in a recent court case, the court ruled that the Medicine Board must give notice to a hospital or clinic when it plans to enter and review records. Moreover, peer review records can be obtained only upon subpoena.

The Medicine Board is required to report its findings to the Massachusetts legislature. Consideration is now being given to how the data should be displayed and how adjustments should be calculated so that providers are represented fairly. In turn, the information prepared by the Medicine Board will be directly available to consumers. Organizationally, the Medicine Board is located under the Massachusetts Office of Consumer Affairs.

Administration require hospitals to have risk-management programs. Massachusetts and New York require that hospitals submit incident reports directly to State authorities (see boxes 5-A and 5-B). The Veterans Administration requires that summaries of incidents be collected centrally.

At present, incident reporting and occurrence screens are in widespread use only in hospitals, but conceptually, there is nothing to preclude their use in other health care settings. Massachusetts already requires that certain kinds of incidents be reported by physicians in office practice to the State Medicine Board (243 CMR 3.11 (1987)).[2] To be of use in ambulatory settings, the screening criteria used in hospital inpatient systems would have to be redefined to identify the adverse events that occur in ambulatory settings. The Public Citizen Health Research Group has suggested screening in ambulatory settings, for example, for the misprescribing of antibiotics such as chloramphenicol, which is rarely medically indicated for ambulatory patients and can cause severe adverse reactions (712). The Health Care Financing Administration (HCFA) has developed criteria for

---

[2] Incidents that must be reported by physicians include: 1) "unplanned transfer to a hospital precipitated by an invasive procedure performed in the office"; and 2) "major or permanent impairments of bodily functions or death that are not ordinarily expected as foreseeable results of the patient's condition or of appropriately selected and administered treatment" (243 CMR 3.11 (1987)).

Box 5-B.—Mandatory Incident Reporting in New York

Since October 1985, first under the general authority of the Commissioner of Health and later in 1986 under statutory authority of the New York Public Health Law, hospitals in New York have been required to report incidents to the State Department of Health within 24 hours of the incidents' occurrence. Hospitals are further required to investigate the incidents and file copies of their reports with the State. The Public Health Law exempts hospital incident reports from disclosure under the Freedom of Information Law and from civil litigation disclosure proceedings. However, the State Department of Health can release summary statistics, as well as statements of deficiencies generated as a result of departmental investigations (592).

Incidents that must be reported in New York include the following:

- patients' deaths or impairments of bodily functions in circumstances other than those related to the natural course of illness, disease, or proper treatment in accordance with generally accepted medical standards;
- fires in the facility that disrupt the provision of patient care services or cause harm to patients or staff;
- equipment malfunction during treatment or diagnosis of a patient that did or could have adversely affected a patient or health facility personnel;
- poisoning occurring within the facility;
- strikes by facility staff;
- disasters or other emergency situations external to the hospital environment that affect health facility operations; and
- termination of any services vital to the continued safe operation of the health facility or to the health and safety of its patients and personnel (591).

Guidelines provide examples of incidents that would fit into the first category, but hospitals still have considerable leeway in interpreting the regulations. Statewide, there were 19 reported incidents per 100,000 patient days in 1986, but with wide variations in reported incidents among hospitals. The Department of Health suspects that this is largely a function of underreporting.

In March 1987, the State Department of Health released the first annual report on the hospital incident reporting system (593). Patient falls accounted for the greatest number of reported incidents (35 percent), but the second highest category of incidents was those related to a treatment or procedure (21 percent including 109 patient deaths). Summary statistics are reported on a statewide, area, and hospital-specific (but not hospital-identified) basis. A stated goal of reporting these statistics is to increase public awareness and knowledge about hospital care.

screening patients' records for quality problems in hospital outpatient departments, home health agencies, and skilled nursing facilities (652); however, these will not be used for reviewing the ambulatory care received by Medicare beneficiaries until 1989. For occurrence screens, as for some of the other potential indicators of quality of care examined in this report, considerable further research is needed if the intention is to use them in nonhospital settings.

Identifying the occurrence of adverse events/incidents is really a problem-oriented approach to quality assessment. Most reporting systems are the inhouse creations of hospitals designed for their own internal needs. Some reporting systems rely on the review of patients' medical records, while others are independent of existing information systems. With such variability in systems for identifying adverse events/incidents and no standardization of the elements/criteria used in the systems (much less of how data should be collected), how can the reliability and validity of the systems as indicators of the quality of care be investigated?

Some of the specific criteria used in existing reporting systems may prove to be reliable and valid indicators of the quality of medical care. Researchers are currently investigating the usefulness in assessing the quality of care of specific patient outcome measures, including rehospitalizations and targeted mortality rates (170,193, 594). To demonstrate the strengths and weaknesses of using specific criteria to assess quality, this chapter examines intensively one criterion that is frequently found in occurrence screens—namely, nosocomial infections. (Another common element in almost all screens—hospital deaths, or some subset of deaths—is analyzed as a potential indicator of quality in ch. 4 of this report.)

A shortcoming of the use of nosocomial infection rates as a quality indicator is that a single indicator may effectively identify quality problems in a specific type of patient or clinical service but not address problems in other areas of medical care; very poor-quality care may go unregistered. A major strength of existing hospital screening systems may well be the use of multiple criteria to identify problems. On the other hand, multiple variables complicate analysis, even under ideal research conditions. Where relevant research related to occurrence screens has been done, this chapter notes it.

The remainder of the chapter is organized as follows. First, occurrence screens that might be considered standardized because they have been developed at the State or national level are described. Then, the reliability, validity, and feasibility of using either nosocomial infections or "standard" occurrence screens as indicators of the quality of care are examined. Finally, conclusions are stated, and the policy implications of using adverse events as indicators of the quality of care in hospitals are explored.

STATE- AND NATIONAL-LEVEL OCCURRENCE SCREENS

In the vast majority of cases, hospitals design and implement their own screening systems for adverse events. Under development or already in place, however, are a number of national and State-level activities that use the same general methods and approach. In the private sector, for example, the Maryland Hospital Association has undertaken a project to find a limited number of data elements (clinical indicators) that could be commonly defined and would permit meaningful comparisons among hospitals for the purpose of assessing quality. Nine indicators were tested in pilot Maryland hospitals beginning in 1985, and today, following deletions, additions, and revisions of various indicators, the study is being conducted in more than 40 voluntarily participating hospitals. The indicators being studied include nosocomial infections, surgical wound infections, autopsy rates, newborn deaths, perioperative deaths, cesarean sections, hospital readmissions, unplanned admissions following ambulatory surgery, intensive care unit readmissions, and unscheduled returns to the operating room (607).

The State of Pennsylvania's Health Care Cost Containment Council collects data on two elements that are usually considered occurrences, nosocomial infections and hospital readmissions (484). Because the data are collected on every hospital patient discharged, adverse events can be linked to specific physicians and services. Moreover, for every hospitalized patient, Pennsylvania hospitals are required to submit to the State Council an indicator of the severity of illness (Medis-Groups methodology) along with other more standard discharge abstract information. The Pennsylvania reporting system is currently being implemented, and published statistics that include patient severity of illness adjustments are not expected before 1990. Other States have demonstrated interest in similar reporting systems. Colorado, for example, has issued regulations effective January 1989 that require reporting patient severity of illness levels as part of required hospital discharge abstracting systems (140).

JCAHO expects to expand its accreditation activities to include the use of clinical indicators to screen hospital cases for quality problems (324). Three JCAHO task forces, working on obstetrical, anesthesia-related, and hospitalwide clinical indicators, have identified structure, process, and outcome clinical criteria that are currently being tested as screens for quality problems in pilot hospitals. The hospitalwide indicators being evaluated are shown in table 5-2. Also shown in that table are the most important patient risk factors or covariates that might also influence outcomes. JCAHO is continuing to develop indicators for a variety of clinical areas, but use of the clinical

Table 5-2.—JCAHO Hospitalwide Clinical Indicators Being Evaluated as Screens for Hospital Quality Problems

HOSPITALWIDE CLINICAL INDICATORS BEING EVALUATED

1. Unplanned readmissions to a hospital shortly after inpatient surgery
2. Unplanned admissions to a hospital shortly after outpatient surgery or specified procedures
3. Development or worsening of pressure ulcers (decubiti)
4. Development of wound infections after clean or clean-contaminated surgical procedures
5. Development of pneumonia in patients treated in special care units
6. Development of infections related to the use of intravascular devices in special care units
7. Proper timing of antibiotic prophylaxis for specified surgical procedures
8. Appropriate use of blood culture sensitivities in the treatment of bacterial sepsis
9. Development of complications associated with suboptimal methods of administration and monitoring of specified medications
10. Commission of important medication errors resulting in death or major morbidity
11. Mortality of patients with specified medical conditions either during hospitalization or within 30 days of admission if death occurs at another institution to which the patient was transferred
12. Mortality of patients after specified surgical procedures either during hospitalization or within 30 days of admission if death occurs at another institution to which the patient was transferred
13. Mortality among patients treated in the hospital for injuries sustained immediately prior to treatment when death occurs within 30 days of injury or during a hospitalization that was precipitated by the occurrence of the injury

Supplemental information collected

Patient risk factors (covariates) that might influence outcomes

- Age
- Sex
- Height and weight
- Braden Risk Scale on admission to hospital and special care units
- Glasgow Coma Score on admission to hospital and special care units
- Trauma Score of patients prior to treatment for injuries
- Diagnoses on admission to hospital, immediately prior to operation or specified procedure, and on admission to special care units (6 digit ICD-9-CM)
- Types of surgical or other specified procedures, if any (4 digit ICD-9-CM)
- Nature of surgical or other specified procedures, if any (scheduled, urgent, or unscheduled)
- Type and site of intravascular devices used in special care unit
- Selected chronic medications on admission to hospital
- Selected laboratory values on admission to hospital, immediately prior to operation or specified procedure, and on admission to special care units
- Temperature, pulse, respiration, and systolic and diastolic blood pressure on admission to hospital, immediately prior to operation or specified procedure, and on admission to special care units

Other information

- For patient admitted after outpatient procedure: stated reason for admission
- Insertion of drains during clean and clean-contaminated surgery
- Patient with endotracheal tube or tracheostomy in special care unit
- Use of nasogastric tube in special care unit

SOURCE: Joint Commission on the Accreditation of Healthcare Organizations, *National Invitational Forum on Clinical Indicators* (Chicago, IL: November 1987).

indicators as part of the accreditation process is not expected to be fully implemented until 1990 at the earliest.

The U.S. Department of Defense screens about 10 percent of all discharges from its 167 hospitals using exhaustive process and outcome clinical criteria that were developed by consensus panels of experts (447). This screening takes place under the Department of Defense Civilian External Peer Review Program. All cases involving 1 of 34 specific diagnoses or 14 problems are sampled, and the patients' medical records specially abstracted by medical record technicians. The abstracted information is computerized, and the screen of clinical criteria then applied. Cases failing the computer screen (about 10 to 20 percent fail) are reviewed by physicians.

The adverse event screening program that has had the most far-reaching impact to date is HCFA's "generic quality screen," which is used to screen hospitalized Medicare patients for quality problems (see table 5-3). Since July 1986, utilization and quality control peer review organizations (PROs) have been required to apply the generic screens to every case they review (about one-fourth of all Medicare discharges). Nurse reviewers examine patients' medical records, and if a screen is failed, the medical record is referred

Table 5-3.—HCFA's Generic Quality Screens[a]

*1. **Adequacy of discharge planning**
   No documented plan for appropriate followup care or discharge planning as necessary, with consideration of physical, emotional, and mental status/needs at the time of discharge.

2. **Medical stability of the patient at discharge**
   a. Blood pressure on day before or day of discharge
      systolic—less than 85 or greater than 180
      diastolic—less than 50 or greater than 110
   b. Temperature on day before or day of discharge greater than 101° F oral (rectal 102° F)
   c. Pulse less than 50 (or 45 if the patient is on a beta blocker), or greater than 120 within 24 hours of discharge
   d. Abnormal results of diagnostic services which are not addressed or explained in the medical record
   e. Intravenous fluids or drugs on the day of discharge (excludes KVOs, antibiotics, chemotherapy, or total parenteral nutrition)
   f. Purulent or bloody drainage of postoperative wound within 24 hours prior to discharge

3. **Deaths**
   a. During or following elective surgery
   b. Following return to intensive care unit, coronary care or special care unit within 24 hours of being transferred out
   c. Other unexpected death

*4. **Nosocomial infections**
   a. Temperature increase of more than 2° F more than 72 hours from admission
   b. Indication of an infection following an invasive procedure (e.g., suctioning, catheter insertion, tube feeding, surgery)

5. **Unscheduled return to surgery within same admission for same condition as previous surgery or to correct operative problem** (exclude "staged" procedures)

6. **Trauma suffered in the hospital**
   a. Unplanned removal or repair of a normal organ (i.e., removal or repair not addressed in operative consent)
   *b. Fall with injury or untoward effect (including but not limited to fracture, dislocation, concussion, laceration, etc.)
   c. Life-threatening complications of anesthesia
   d. Life-threatening transfusion error or reaction
   e. Hospital acquired decubitus ulcer
   f. Care resulting in serious or life-threatening complications, not related to admitting signs and symptoms, including but not limited to the neurological, endocrine, cardiovascular, renal or respiratory body systems (e.g., resulting in dialysis, unplanned transfer to special care unit, lengthened hospital stay)
   g. Major adverse drug reaction or medication error with serious potential for harm or resulting in special measures to correct (e.g., intubation, cardiopulmonary resuscitation, gastric lavage) including but not limited to the following:
      i. Incorrect antibiotic ordered by the physician (e.g., inconsistent with diagnostic studies or the patient's history of drug allergy)
      ii. No diagnostic studies to confirm which drug is correct to administer
      iii. Serum drug levels not performed as needed
      iv. Diagnostic studies or other measures for side effects not performed as needed (e.g., BUN, creatinine, intake and output)

[a] For entries marked with an asterisk in this table, the PRO reviewer is to record the failure of the screen, but need not refer to physician reviewer.

SOURCE: U.S. Department of Health and Human Services, Health Care Financing Administration, Health Standards and Quality Bureau, *1986-1988 PRO Scope of Work*, Baltimore, MD, Nov. 4, 1985.

to a physician advisor for further review. Only the physician advisor can declare a case a "quality problem." On the basis of this information, the PROs build provider profiles for their own internal use; they also take corrective actions ranging from education to intensified review, and ultimately to sanctions (see ch. 6). Appendix D provides a full description of the PROs' review procedures and responsibilities.

RELIABILITY OF THE INDICATOR

Nosocomial (Hospital-Acquired) Infections

As noted earlier in this chapter, OTA chose an adverse outcome used in almost all existing occurrence and generic screens for indepth review—namely, nosocomial infections. One reason for selecting nosocomial, or hospital-acquired, infections is that such infections are quite prevalent. Six percent of all U.S. hospitalizations are complicated by nosocomial infections, amounting to more than four million nosocomial infections per year (269). Of course, infections may also be acquired in the community prior to admission to the hospital. Nosocomial infections are defined as infections that are not known to be present or incubating at the time of admission. The most common nosocomial infections are urinary tract infections (42 percent), followed by surgical wound infections (24 percent), pneumonia (10 percent), and infections of the bloodstream (bacteremia) (5 percent). These four types of infections account for about 80 percent of all nosocomial infections. Almost three-quarters of nosocomial infections occur among patients undergoing surgery (273).

The most difficult obstacle to the reliable measurement of nosocomial infections is the lack of standardized case finding. Reliable measurement of infections requires that trained surveillance personnel search actively for cases using standardized clinical definitions of infections (269). No system of routine data collection is completely sensitive in identifying nosocomial infections, and the surveillance techniques that are used in case finding in various hospitals differ fundamentally (616). The likelihood that nosocomial infections will be clearly recorded in a patient's medical record and/or coded on a hospital discharge abstract varies widely by hospital, but relying on written diagnoses is generally an inaccurate method of determining infection rates (232). One study in a university hospital found, for example, that 43 percent of nosocomial infections were not coded in the hospital discharge abstract (409).

A study sponsored by the Centers for Disease Control (CDC) showed that reliable measurement of, and changes in, nosocomial infection rates at various sites are possible in a large-scale data collection effort that relies on medical record review (275). The Study on the Efficacy of Nosocomial Infection Control (SENIC) Project evaluated the efficacy of the infection surveillance and control programs established between 1970 and 1975-76 in a representative sample of U.S. hospitals. Incidence rates of nosocomial infections in four sites (urinary tract, surgical wound, lower respiratory tract, and bloodstream) were determined from a random sample of medical records in each of the 2 years in 338 hospitals stratified by size, teaching status, and infection control activity. To measure nosocomial infection rates reliably, CDC devised a standardized method of making diagnoses via retrospective review of patients' medical records, and it validated the method's accuracy through a series of pilot studies. Nonphysician CDC reviewers, who underwent careful training and infield monitoring, abstracted relevant data, recorded them on standardized forms, and applied a set of standardized algorithms to arrive at the infection diagnoses.

The retrospective chart review method used in the SENIC Project (by nonphysicians following a standardized procedure) compared favorably (average sensitivity of 0.74) with the "gold standard" method of physician-epidemiologists supervising intensive prospective data collection teams (230,275). As measured against this standard, physician self-reporting forms were least sensitive (0.14 to 0.34) in finding cases of nosocomial infection, and clinical surveillance for evidence of fever, antibiotic use, or both were only moderately sensitive (0.47 to 0.59) (230).

Because the recognition of infections depends in part on physicians' propensity for ordering the cultures and chest X-rays that confirm the presence of infection, the SENIC Project also analyzed the use of these diagnostic tests in the sample hospitals (270). Generally, the researchers found an increase over time in the use of diagnostic tests, and the increased use of these tests was associated with increased recognition of infectious diseases. More importantly, despite clinical agreement on the efficacy of these diagnostic tests, hospitals differed significantly in diagnostic medical practices. Hospitals with high rates of culturing, working up fevers, and obtaining chest X-rays showed higher *observed* rates of nosocomial infections.

This finding presents an additional measurement problem that cannot be resolved through better or standardized data collection efforts (270). If nosocomial infection rates were used as indicators of quality in cross-hospital comparisons, those hospitals that were effectively identifying nosocomial infections through appropriate testing could be penalized. Because no diagnostic testing is necessary to confirm the presence of surgical wound infections, a possible solution would be to compare infection rates only for this subset of nosocomial infections (270,305).

HCFA's Generic Quality Screens

As shown in table 5-3, HCFA's generic quality screens apply two criteria related to nosocomial infections (item 4): "a) temperature increase of more than 2 degrees more than 72 hours from admission; and b) indication of an infection following an invasive procedure." Depending on how individual PROs interpret and use the nosocomial infection screens, results could vary greatly. Nurse reviewers searching for "indications of an infection," for example, could either rigorously review all laboratory records, progress notes, and nursing notes or simply look for documentation of antibiotic use or specific laboratory test results. According to an initial report from HCFA on the use of the generic quality screen by PROs, more of the discharges reviewed failed the nosocomial infection screen (5 percent) than any other (except medical stability at discharge), but fewer than 15 percent of these cases upon further review by a physician advisor actually had a significant medical problem (653).

The physician advisor must decide which of the discharges that have failed the nosocomial infection screen constitute actual "quality problems." There are no guidelines on how clinically to ascertain a quality problem; the judgment is primarily subjective. Thus, at present, the same case that is considered a problem in one PRO (or by one physician advisor) might be discounted by another PRO. In some PROs, for example, the physician advisors were not counting nosocomial infections as quality problems if the infections were treated appropriately (487). Recently revised guidelines on the application of the generic quality screens clarify that nosocomial infections should be counted regardless of therapy (652). Nonetheless, there is obviously a severe reliability problem with HCFA's generic quality screen that results from the subjective nature of the physician advisor's audit.

Only summary data on the generic quality screens (neither hospitals nor physicians are identified) are forwarded by the PROs to HCFA. Data reported to HCFA for the first year during which the generic screens were used showed wide variation in the incidence of screen failures and of confirmed quality problems across PROs (660). In several PROs, fewer than 5 percent of cases failed any screen; in other PROs, more than 40 percent failed. In cases of screen failures, the percentage of confirmed quality problems ranged from zero to 100 percent.

To ameliorate substantial reliability problems, the so-called SuperPRO, an independent contractor, is charged to re-review a sample of each PRO's cases to validate the determinations of nurse reviewers and physician advisors. In its first review of the application of the generic quality screens in 45 PROs, the SuperPRO found 8.9 percent of sample cases with quality problems v. only 3.8 percent reported by the PROs (654).

In response to critiques, HCFA has revised the generic quality screens for the PROs' third round of contracts, which will probably begin in early 1989 (652). The revised generic quality screens have several changes (see app. D). In the future, for example, nurse reviewers will flag a case as a nosocomial infection only if two or more indications listed in new HCFA guidelines are present in a patient's chart.[3] In addition, all PROs have

Photo credit: George Washington Medical Center

Rates of surgical wound infections are potentially valid indicators of the quality of care in hospitals.

---

[3] Indicators of a nosocomial infection include: temperature elevation of 101 degrees Fahrenheit or greater; elevated white blood count and/or left shift; isolation of organism from body fluids or specimens; appropriate radiographic imaging abnormalities; purulent drainage; heat, redness, focal tenderness and/or pain; pyuria, dysuria; and productive cough (652).

been issued the CDC guidelines for the surveillance of nosocomial infections.

These steps are likely to improve the reliability of HCFA's generic quality screens over the next several years. Unlike the CDC personnel in the SENIC Project, however, PRO reviewers do not receive intensive training in the use of the guidelines, nor do they use diagnostic algorithms. Moreover, the audit by physician advisors of cases that fail the screens is largely subjective. The SuperPRO has now started to analyze the reliability of PRO results for individual generic screen criteria. Depending on the findings of these analyses, further revisions of HCFA's generic quality screens may be necessary in the future.

VALIDITY OF THE INDICATOR

Nosocomial (Hospital-Acquired) Infections

Numerous studies link nosocomial infections to lengthened hospitalization, morbidity, and/or mortality (160,233,251,261,263,493,536,587,698). A prospective study of patients with indwelling bladder catheters in a teaching hospital, for example, found the development of urinary tract infections among these patients to be associated with a threefold increase in mortality (493). One analyst estimates that more than $2.8 billion in excess hospital charges are generated each year because of nosocomial infections (182). Because of the empirical association of nosocomial infections with adverse outcomes for patients, nosocomial infections have high face validity as an indicator of the quality of medical care.

Although the relationship between nosocomial infections and poor patient outcomes is well established, the link between inadequate/poor hospital care and the onset of infection is less clear. The fact that an infection is acquired in the hospital does not mean that it is *caused* by the hospital or by the poor quality of its practitioners. No available studies have examined or compared nosocomial infection rates in hospitals explicitly to examine the quality of providers. Numerous studies have published institutional nosocomial infection rates, however, as part of investigations of effective interventions, changes in rates over time, or the health and cost implications of hospital-acquired infections.

A review of the literature through 1975 identified 24 studies that published survey data on nosocomial infections in hospital populations (230). The prevalence of nosocomial infections in the hospital populations in these data ranged from 4.5 to 15.5 percent, and the incidence of such infections (infections per 100 discharges) varied from 3.1 to 14.1 percent. Community hospitals had lower reported nosocomial infection rates than referral, municipal, or chronic disease hospitals.

Comparisons of data from these studies tell little about the quality of care in the hospitals surveyed because, aside from measurement problems, the data are not adjusted for the hospitals' case mix or patients' severity of illness. Although most of the studies report nosocomial infections by site of infection, by service, and by procedure, the samples are too small to allow adequate stratification of the patient populations. Researchers attempting to calculate the impact of nosocomial infections on morbidity and costs usually compensate for confounding variables by matching infected patients with comparison subjects on as many attributes as possible. Although the results may be valid for the institution studied, it is very difficult to compare study results across institutions, even for seemingly similar subgroups of patients (e.g., all surgical patients or all patients with the same primary diagnosis). The authors of the literature review just mentioned attempted to compare the results of their matched subject study at Boston City Hospital with three other epidemiologic reports. Inconsistencies in results were attributed to possible further confounding variables among the patient populations (231).

The risk of acquiring a nosocomial infection is related to a number of factors in addition to the quality of providers. The likelihood of an infection's occurring and its outcome depends more on patient susceptibilities than on the presence of the organism (49). Patients' underlying diseases, medical procedures, severity of illness at admission, hospital service, age, sex, race, and urgency of admission have all been found to be significant risk factors for nosocomial infection (96,232).

Understanding and adequately adjusting for such risk factors are critical to the use of nosocomial infections as a valid indicator of the quality of care. Moreover, the necessary adjustment factors for nosocomial infections may be different from those used to compare mortality statistics or other quality indicators. For example, one study, which compared urinary tract infections in small hospitals (under 75 beds) with infections in a large, teaching hospital, observed that the higher prevalence rate in the teaching hospital was due to the increased use of indwelling bladder catheters (95). With even a rudimentary understanding of case mix, it is not surprising that community hospitals have lower rates of nosocomial infections than teaching and municipal hospitals.

The SENIC Project provides valuable information, because the researchers attempted to control for patient risk and other intervening factors in their investigation of the efficacy of infection-control programs. Using the large SENIC data base, the researchers determined estimates of the frequency of nosocomial infection by selected characteristics of patients (273).[4] Hospital-related characteristics were controlled by using American Hospital Association survey data as proxies for changes in hospitals that could not be measured (272). And finally, differences in physicians' diagnostic practices (their propensities for ordering tests) were controlled by defining hospital-specific measures for use in analyses (272).

Because of confidentiality provisions, the SENIC Project data cannot be analyzed by hospital. Nevertheless, the research helps to validate nosocomial infection rates as quality indicators in several ways. First, the SENIC Project researchers have measured and quantified the patient risks and other variables that contribute to the outcome of hospital-acquired infection. This information could be used in further research to allow comparisons among hospital populations. The large, statistically valid data base developed in the SENIC Project could permit "norms" for nosocomial infection rates to be established by patient risk category.[5]

Second, in concluding that one-third of all nosocomial infections could be prevented through surveillance and control programs, the SENIC Project demonstrates the potential of nosocomial infection rates to serve as an indicator of the quality of care across hospitals (272). One infection-control program shown to be efficacious was the systematic feedback of surgical wound infection rates to the practicing surgeons. (In combination with an ongoing surveillance and control program, this program led to a 19-percent decrease in surgical wound infections.) Thus, changes in physicians' behavior, or the process of care, are associated with changes in nosocomial infection rates. Moreover, the extent to which hospitals establish and maintain effective infection control programs is an aspect of their quality of care.

There is no evidence that nosocomial infection rates are correlated to the general quality of health care institutions (external validity). In fact, there are well-defined inpatient groups who have very little risk of acquiring nosocomial infections, for example, pediatric, psychiatric, and rehabilitation patients (274). Nosocomial infections would not be valid indicators of the quality of care received by such patients.

Occurrence Screens

Most occurrence screens are based on the criteria established in the California Medical Insurance Feasibility Study, which reviewed a large sample of 1974 California hospital records. The study sought to identify potentially compensable events or medically caused patient disabilities. For the purposes of this OTA assessment, the potentially compensable events identified in the study are synonymous with adverse events caused by poor-quality care. Investigators in the California study sampled hospital charts by service from a group of 23 hospitals stratified by size, ownership, and teaching status. Of the more than 20,000 charts reviewed by medical record auditors, approximately 50 percent failed the screens. The study investigators (all physicians) reviewed these records and concluded that 11 percent of those failing the screens constituted potentially compensable events (or 5.5 percent of all records reviewed) (432).

---

[4] Risks were significantly related to age, sex, service, duration of total and of preoperative hospitalization, presence of previous infection, types of underlying illnesses and operations, duration of surgery, and treatment with urinary catheters, continuous ventilatory support, or immunosuppressive medications (273).

[5] CDC has another ongoing data collection system, the National Nosocomial Infections Surveillance System, that collects nosocomial infection rates from 85 volunteer hospitals. CDC is using these more recent data to develop risk indices by diagnosis-related groups and for surgical, critical care, and neonatal intensive care patients (305).

The California Medical Insurance Feasibility Study validated its 20 screening criteria as part of a controlled two-step screening and audit process for determining the incidence of potentially compensable events. It usefully identified potentially compensable events by medical specialty, location (e.g., 72 percent of the potentially compensable events occurred in the operating room), diagnosis and procedure, and by selected characteristics of patients. However, the study did not validate the screening criteria (by themselves) as quality indicators. In fact, on the basis of the published data, it is not possible to calculate the sensitivity or the specificity of the screening criteria in identifying either potentially compensable events or adverse events (potentially compensable events are a subset of adverse events that are medically caused). There is insufficient information about the patients' medical charts that passed the screens to determine these values. Moreover, of the records in the study that failed the screens, 81 percent were eliminated by the investigators because no medically or patient-caused disabilities were found upon further examination of the records. This high percentage indicates a substantial false-positive problem, whether the goal of the screens is identification of adverse events or identification of potentially compensable events. The two-step screening and audit process may be a valid and effective, yet very inefficient, method of identifying poor-quality care.

The California study did not examine the effectiveness of individual criteria in screening for potentially compensable events. Moreover, the determination by the physician investigators of whether a potentially compensable event occurred was largely subjective (as is also true in the PRO program). The subjectivity of such assessments is a critical factor in the reliability of audit when more than just a few investigators are involved.

Research commissioned by New York State under recent medical malpractice reform legislation will update the results of the California study and help to ascertain the validity of occurrence screens. As part of a comprehensive study to find which patients suffered injuries in the course of their hospital treatment and which of these injuries were produced as a result of substandard treatment, the Harvard Medical Practice Study Group is reviewing the medical records of 30,000 patients hospitalized in New York in 1984. These records are being reviewed by medical record administrators using 17 screens derived from the 1974 California Medical Insurance Feasibility Study.[6] The medical records that fail the screens are then subjected to further review by physicians to confirm the adverse event, to estimate the probability of causation, and finally to estimate the probability of negligence (283).

The results of the Harvard study commissioned by New York State could validate the relationship of the screening criteria (outcome measures) to poor-quality care (the process of medical management) if the data are directed to that purpose. The Harvard study may reaffirm the finding of the California study that occurrence screening as part of a two-step process involving screening and subsequent audit is a valid approach to quality assessment. The relationship of the screening criteria to the universe of adverse events or poor-quality care in hospitals, however, will not be resolved adequately by the Harvard study. Because the medical records that do not fail the screens are not examined in depth, the true denominator number of adverse events remains unknown. The full-scale study began in mid-1987, and results are expected in early 1989.

---

[6] The California screens have been modified by the deletion of four criteria ("unplanned removal of an organ or part of an organ during an operative procedure," "wound infection on last full day prior to or day of discharge," "discharge with indwelling urinary catheter," and "parental analgesics last full day prior to discharge") and the addition of one criterion ("obstetric mishap or complication of abortion, labor, or delivery") (283).

The SuperPRO has evaluated the accuracy of HCFA's generic quality screens in finding quality problems. In a special study, the SuperPRO reviewed a sample of medical records from nine PROs for the period August 1986 through January 1987 (444). Just as the PROs do, the SuperPRO's nurse reviewers applied HCFA's generic screens and referred cases that failed to physician reviewers for determination of quality problems. In addition, the SuperPRO calculated how many false negatives the screening process yielded by sampling the records that had passed the generic screens. These records were re-reviewed by a physician to determine if there were quality problems.

The SuperPRO concluded that HCFA's generic screening process had a sensitivity (i.e., ability to identify cases with quality problems) of 49 percent and a specificity (i.e., ability to exclude cases without quality problems) of 73 percent (444). A sensitivity of less than 50 percent means the screening process was no better at detecting quality problems than chance. Because a small sample size (100 records) was used by the SuperPRO to determine the false negatives, the sensitivity finding may have some degree of error and may actually range between 37 and 70 percent. In any event, the SuperPRO researchers concluded the quality problems that were found through HCFA's generic screening process were more serious than the quality problems missed by the process.

The SuperPRO also evaluated individual screening criteria used in HCFA's generic screen, especially those criteria thought to be responsible for substantial numbers of false positives. The study recommended dropping several screening criteria (including one related to nosocomial infections) and modifying several others. HCFA's revisions of the generic quality screen for the 1988-90 PRO contract cycle were a response to these recommendations (see app. D).

The SuperPRO study is useful insofar as it relates to the validity and effectiveness of individual criteria, but it also has several shortcomings. The study's sample of Medicare cases, for example, is not a random sample; it is probably weighted toward problem cases. In addition to reviewing a mandatory random 3-percent sample of hospital discharges, PROs review cases based on a number of negotiated objectives. In selecting its re-review sample, the SuperPRO did not distinguish among the types of cases reviewed by the PROs. Moreover, the small sample size used in the special SuperPRO study does not permit reliable estimates of the validity of the screening process. The SuperPRO may undertake a larger analysis in the future.

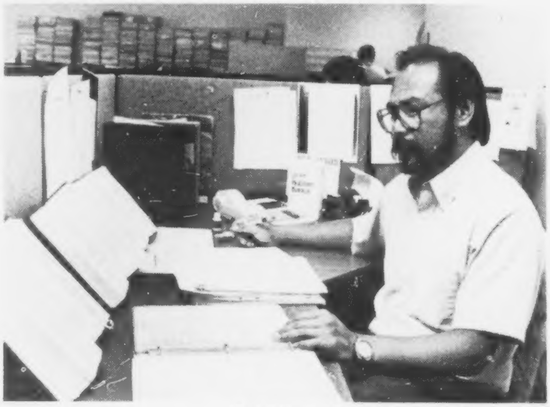


Photo credit: California Medical Review, Inc., Chindy Charles, Photographer

HCFA's generic quality screens, which PRO reviewers apply to all Medicare cases, have not been validated empirically.

FEASIBILITY OF USING THE INDICATOR

Nosocomial (Hospital-Acquired) Infections

The feasibility of obtaining nosocomial infection rates by standardizing data collection methods in all hospitals and maintaining reliability over time is questionable (269,305). Relying on coded diagnoses from hospital discharge abstract systems would be an unreliable method of establishing infection rates across hospitals. At a minimum, thorough medical record review by trained personnel is essential for finding cases of nosocomial infections. The PRO audit process involves such thorough chart review by nurse reviewers with followup by physician advisors.

An alternative to medical record review would be to establish new channels to obtain more reliable data. Currently, for example, all hospitals are required to have designated infection-control personnel and infection-control committees in order to be JCAHO accredited and to be eligible for Medicare and Medicaid reimbursement (424). Infection-control officers, usually nurse and sometimes physician epidemiologists, use ongoing surveillance techniques to find cases of nosocomial infections. If infection-control officers were required to use the standard definitions and guidelines provided by CDC, the data obtained by these personnel and utilized by the infection control committees could be channeled outside the institution for quality assessment purposes. CDC currently collects such data from approximately 85 volunteer hospitals in its National Nosocomial Infections Surveillance System (305).

Using rates only for selected sites of nosocomial infections, such as the bloodstream and surgical wounds, rather than combined rates of nosocomial infections at all sites, would minimize the measurement problem created by differing physician diagnostic practices. Bloodstream infections, which require only one verifying laboratory culture, have been suggested as one type of nosocomial infection for which reliable statistics could be gathered (305,699). Surgical wound infections do not require laboratory verification, although an impartial view of the wound in the operating room is necessary to determine the degree of contamination before and during the operation. Moreover, research has progressed furthest in understanding confounding patient risks for surgical wound infections. Data from the SENIC Project were analyzed using multiple logistic regression techniques (271). The researchers concluded that four risk factors predict a patient's probability of getting a surgical wound infection twice as well as the traditional classification of wound contamination alone: abdominal operation, operation lasting more than 2 hours, contaminated or dirty-infected operation, and three or more underlying diagnoses.

Occurrence Screens and Incident Reporting

The use of occurrence screens and incident reporting by hospitals is widespread. The general availability of such systems was a primary reason for OTA's decision to study adverse events as a potential indicator of the quality of care. To the extent that occurrence screen and incident reporting systems are already in place, the additional costs of supplying information on adverse events to consumers could be minor as compared with costs of supplying information on other quality indicators. Moreover, poor patient outcomes are readily understandable by consumers and associated in the public mind with the quality of care.

Regulators are increasingly turning to occurrence screen and incident reporting systems to accomplish their goals in quality assurance. New York, and more recently Massachusetts, are collecting incident reports and, in turn, making selected information publicly available. Pennsylvania is implementing a statewide hospital discharge abstract system that includes information on the patient's severity of illness at admission and on several data elements normally considered occurrences or adverse outcomes. A primary purpose of Pennsylvania's data system is to inform the public about health care costs and quality. Several other States, including Colorado and Iowa, are pursuing approaches similar to Pennsylvania's. Thus, a number of State-level systems either already are, or soon will be, using statistics on adverse outcomes to inform consumers about the quality of hospitals.

On the national level, hospital-specific data generated by the PROs through the application of HCFA's generic quality screens are available to the public upon request to a PRO, subject only to hospital notification at least 30 days before disclosure (42 CFR 476.120,476.105 (1987)). Consumers can request information by hospital on screen failures, on quality problems identified during audit, or on both. As far as HCFA is aware, no such requests of PROs have been made to date (487). The Public Citizen Health Research Group contends that at least one PRO has refused to make similar types of outcome data available to public requestors even though it is legally required to do so (713).

To the extent that incidents and occurrences are reported through inhouse systems (without independent audit by outside quality assessors), hospitals have plentiful opportunities to underreport or to "game" the results. The congressional General Accounting Office investigated the Veterans Administration's incident reporting system and found that 86 percent of the incidents occurring in a sample of cases were unreported (624). The disincentives for hospitals to report adverse events are obvious: possible malpractice litigation or other disciplinary action and recognition as a poor-quality provider. New York State relies on several other systems it has in place, including State accreditation surveys, patient complaints, and special studies, to verify the accuracy of incident reporting by hospitals. Nonetheless, despite such possible cross-checks on hospitals, the reliance of most occurrence and incident reporting systems on self-reporting is a major shortcoming with regard to their use as quality indicators.

CONCLUSIONS AND POLICY IMPLICATIONS

As this chapter has shown, a number of systems for reporting adverse events in hospitals are in place and either are, or could be, used to inform consumers about the quality of care in these institutions. Unfortunately, however, none of these systems have been adequately validated. Data on the number of screens failed or the overall number of self-reported incidents alone are clearly not valid quality indicators and would be meaningless and misleading if used to compare hospitals. The screens in place were not designed to measure quality directly, and substantial proportions of cases that fail the screens, variably across institutions, turn out on further review to be false positives. Moreover, incident or occurrence reporting systems that rely solely on self-reports are unreliable sources of information.

On the other hand, several systems that employ a two-stage process of screening and intensive auditing have been partially validated for quality assessment. Access by consumers to the end results of these assessments has great potential. Two primary unresolved problems that need to be addressed through further research are the extent to which these systems do not identify quality problems that actually exist and the subjective nature of professional audits.

Some of this research is already underway or could be easily undertaken. New York State, in its Harvard study, is investigating a screening and audit method of identifying problem care. JCAHO is studying clinical indicators that will operate at the hospital service level and can be analyzed using covariates of patient risk. Various other efforts, for example, by the Maryland Hospital Association and the Pennsylvania Health Care Cost Containment Council, are underway to verify, define, and/or standardize useful adverse outcome measures for quality assessment.

Further research on the validity of HCFA's generic quality screens for quality assessment is also merited. The screens were developed primarily by professional consensus, and the screen elements have not been validated in empirical studies. HCFA could provide leadership on such research. HCFA's generic quality screens are applied to more hospitalization reviews than any other standardized occurrence screen, and potentially, the results of these reviews could be made easily accessible to the public.

Because all the systems described in this chapter are very new (virtually all have been started during the past several years or are still being implemented), many independent research initiatives are probably useful and appropriate. Pursuing many similar approaches has the potential benefit of developing a wholly new, more effective and efficient system. The rush of State officials and others to implement some kind of quality assessment system means the results of research need to be shared in as timely a fashion as possible. For those systems where new data collection systems are required, a major concern is that different measures or definitions will be used in various systems and the ability to link systems in the future will be lost. Thought should be given now to such long-term needs of uniform reporting and linkage among various State systems.

Another concern is that, because some occurrence screen and incident reporting systems are in operation and the data can be accessed, statistics about adverse events might be released prematurely and misinform the public. None of the systems now in place is specifically designed to provide comparative information about the quality of hospitals. Regulatory agencies employ the systems to target their review or investigations. The potential misuse of information about adverse events in hospitals gives added impetus to the need for research on the validity and reliability of this indicator.

Chapter 6

Disciplinary Actions, Sanctions, and Malpractice Compensation

CONTENTS

Page
Introduction 121
Disciplinary Actions by State Medical Boards 122
  Reliability of the Indicator 123
  Validity of the Indicator 123
  Feasibility of Using the Indicator 126
Sanctions Recommended by Peer Review Organizations and Imposed by HHS 126
  Reliability of the Indicator 127
  Validity of the Indicator 130
  Feasibility of Using the Indicator 132
Malpractice Compensation 133
  Reliability of the Indicator 134
  Validity of the Indicator 134
  Feasibility of Using the Indicator 137
Conclusions and Policy Implications 138
  Disciplinary Actions by State Medical Boards 138
  Sanctions Recommended by PROs and Imposed by HHS 139
  Malpractice Compensation 140
  Combinations of Indicators 141

Figures

Figure *Page*
6-1. Overview of the PRO/HHS Sanction Process for Substantial Violations . 128
6-2. Overview of the PRO/HHS Sanction Process for Gross and Flagrant Violations 128

Tables

Table *Page*
6-1. Physician Disciplinary Actions by State Boards, 1986 124
6-2. Interjudge Consistency in Complex Human Judgments 135

Chapter 6

Disciplinary Actions, Sanctions, and Malpractice Compensation

INTRODUCTION

Federal and State laws and regulations and private sector medical entities have established many methods to discipline and sanction errant members of the medical profession. This chapter evaluates as possible indicators of the quality of medical care three such activities:

- disciplinary actions taken by State medical boards,[1]
- sanctions recommended by utilization and quality control peer review organizations (PROs) and imposed by the U.S. Department of Health and Human Services (HHS), and
- malpractice compensation, particularly court awards.

Disciplinary actions by State medical boards, PRO/HHS sanctions, and malpractice compensation, either separately or in conjunction with each other and other indicators, may have the potential to identify physicians who do not follow accepted standards of care. Those physicians who are disciplined, sanctioned, or successfully sued for malpractice may actually provide substandard care. On the other hand, not all physicians who provide substandard care are disciplined or successfully sued. Studies of avoidable injuries indicate that the universe of avoidable adverse outcomes may be significantly greater than the number of disciplinary actions, sanctions, and malpractice suits (152,595). These studies suggest a large number of poor-quality physicians are not identified or penalized, thereby pointing to the ineffectiveness of existing systems to identify all those individuals providing poor-quality care.

This chapter uses procedures somewhat different from those described in appendix C to evaluate the reliability and validity of disciplinary actions, sanctions, and malpractice compensation as indicators of the quality of care. There are two reasons for modifying the procedures described in appendix C when considering these three indicators. First, the procedures described in appendix C apply to a systematic synthesis of the literature, and studies that examine the causal relationship between any of the three indicators discussed in this chapter and the quality of care are not available. In the absence of research studies, this chapter uses deductive reasoning from the indirect evidence of descriptive information to provide some insight into the reliability and validity of disciplinary actions, sanctions, and malpractice compensation as indicators of quality.

The second reason for modifying the procedures outlined in appendix C is that the three potential indicators discussed in this chapter are essentially legal processes that rely on judgment and have little or no science base.[2] For purposes of this chapter, the term reliability refers to consistency of the decisions made by a legal body (e.g., disciplinary actions taken by State medical boards). The term validity refers to the scope of the decisions made by a legal body and the capacity of the decisions to actually measure quality. Evidence on reliability and validity is derived from examining the structure of the legal bodies, the grounds for taking actions, the procedures used in taking actions, and the types of actions taken. In the case of disciplinary actions by State medical boards and PRO/HHS sanctions, judicial review of the actions is also examined.

A possible confounding issue in OTA's analysis is that the reliability and validity of disciplinary actions, PRO/HHS sanctions, and malpractice compensation as indicators of the quality of medical care depend to a large extent upon peer review.[3] Differences in criteria used by peer physicians, even experts, in making decisions about medical diagnosis and treatment are well documented (71,185). Such differences may have troublesome implications for the reliability and validity of expert peer opinion in disciplinary actions taken by State medical boards, sanctions recommended by PROs and imposed by HHS, and malpractice compensation.

---

[1] In the following discussion, State licensing bodies and State disciplinary bodies will be called State medical boards, although their official titles as well as their organizational loci vary among States.

[2] Reliability and validity, as described in app. C, are concepts used in applied social science and are not traditionally associated with legal systems.

Analyses of the reliability and validity of disciplinary actions, sanctions, and malpractice compensation as indicators of the quality of care are presented below. Also presented are analyses of the feasibility of using each indicator. The final section of this chapter draws conclusions about the current usefulness of the actions, used singly and together, as quality indicators; suggests methods for improving the reliability and validity of the three actions as quality indicators; and discusses current and future means of disseminating information about the three.

DISCIPLINARY ACTIONS BY STATE MEDICAL BOARDS

The legal authority for licensing physicians to practice medicine and for restricting or revoking licenses rests with the States. In most States, the same body that grants licenses to applicants that it has determined are qualified to practice medicine also disciplines physicians who it has decided are unfit to continue practice (32,260). All State medical boards have the authority to revoke or suspend a physician's license. Other disciplinary actions include probation, limitations, fines, reprimands, letters of censure, letters of concern, and collecting costs of proceedings (206). The general grounds for disciplinary actions are unprofessional conduct or professional incompetence (32). The medical practice act of each State mandates specific grounds, such as incorrect drug prescription and substance abuse, for disciplining physicians.

Medical licensure is intended to grant the privilege of practicing medicine to individuals who are of good moral character and are competent to provide safe care to the public (70 Corpus Juris Sec. 19), but it does not ensure continuing competence—an important issue in light of changing medical knowledge and techniques. The purpose of disciplinary actions by State medical boards is to "protect the public against unfit practitioners" (70 Corpus Juris Sec. 35). State medical boards, which historically have been very conservative in censuring physicians (208), have increased their activity in recent years. Disciplinary actions increased from 1,540 in 1984 to 2,108 in 1985 (91) to 2,302 in 1986 (240). Nonetheless, the percentage of practicing physicians disciplined in 1986 (0.50 percent)[4] is significantly less than the 5 to 15 percent of physicians that some authors have hypothesized to be professionally incompetent to practice (169,208). Although the effectiveness of State medical boards in taking disciplinary actions is an important quality concern, the more specific intent of this chapter is to evaluate whether the disciplinary actions taken by State medical boards are good indicators of the quality of care.

Disciplinary actions taken by State medical boards are worth examining as a measure of quality, because they have face validity for average consumers. An average consumer would expect that limiting or withdrawing[5] a physician's license to practice medicine indicates that the physician is professionally incompetent and would be concerned about using the physician for health care.

---

[3] State medical boards use the expert opinion of their physician members to interpret and apply the vague language often found in legislation governing license discipline. Furthermore, "expert" peers testify when physicians are brought up for hearings. The entire sanction process within PROs depends upon peer opinion, from the original identification of a possible violation to succeeding reviews of the violation. Peer review is also an important part of malpractice cases that are heard in court. "Expert" peers testify to the standard of care that can be applied to the case and whether the defendant met the standard.

[4] There were 462,126 physicians providing patient care in 1986 (35).

[5] In most cases, revoking a physician's license prohibits him or her from practicing medicine. There have been well-publicized instances in which physicians whose licenses were revoked in one State continued to provide medical care in other States where they held licenses. Public and private efforts have been working to eliminate this problem.

Reliability of the Indicator

Nationwide consistency of disciplinary actions by State medical boards is not to be expected, because the granting and limitation or withdrawal of medical licenses are State responsibilities. The proportion of physicians who have had their licenses revoked or modified varies greatly among States (see table 6-1). Differences in medical performance, legal impropriety, and inaccuracy of reporting among the States can account for only a small fraction of the variation in the proportion. A greater part of the variation is attributable to differences in State laws and regulation, and, perhaps, the intensity with which State medical boards engage in disciplinary activities (499).

A State medical board's discipline of similar cases may differ because of factors that are not related to the quality of care. Important witnesses sometimes fail to appear, physicians' lawyers vary in expertise, and aggravating and mitigating factors, which are not defined in statute or case law but vary from case to case, must be weighed in disciplinary decisions (389). Consistency in decisions is particularly difficult to achieve in types of cases where physicians disagree about what constitutes acceptable practice. In some States (e.g., Colorado and Connecticut), a threat to consistency is that more than one body is involved in disciplinary activities (206).

In general, the reliability of disciplinary actions as an indicator of quality within a State depends on the individual State. An investigation of 24 States by the Office of the Inspector General of HHS found "inconsistencies in the type of disciplinary actions taken in relation to the charges and even in the meanings of the different types of actions" (361), both among and within States. Whether disciplinary actions in other States are erratic, and if so, to what extent, is not known.

For the most part, the consistency of disciplinary actions taken within a State depends on the precision of the language specifying the grounds for discipline. The more vague the language, the greater the possibility for differing interpretations and applicability. Consistency of such actions is also related to the specific violation, since most States have precise grounds for some violations and ambiguous grounds for others. Most State medical practice acts list specific grounds for infractions dealing with drug prescription and use, fraud, and other violations (280,720). On the other hand, few of the States that specify incompetence in the practice of medicine or substandard practice as grounds for disciplinary actions define incompetence precisely. Illinois' Professions and Occupations Code defines "professional incompetence as manifested by poor standards of care" (111). In the face of such indefiniteness, consistency is difficult, and application of the rule requires a case-by-case interpretation of the applicable standard of practice.

A State medical board's composition and operating style also enter into the consistency of its decisions. Particularly if the grounds for disciplinary actions are vague, a State medical board could be arbitrary and capricious in its adherence to law and regulations and allow extraneous facts, such as the race, religion, or community standing of physicians, to enter into their decisions. In addition, most boards are voluntary and work long hours on difficult issues with little financial reward. Extensive caseloads are common (658), and the medical boards are usually limited in their disciplinary performance by staff and funds (361). As a result, the reliability of their decisions may be compromised.

In addition to taking formal disciplinary actions against physicians, State medical boards take informal disciplinary actions (91). The rationale and procedures for informal actions differ among the States. Boards take several times more informal than formal actions (91). In some States, informal disciplinary actions are taken because of a lack of investigatory resources and the backlog of unheard cases that most boards currently face (658). In other States, informal actions are used as a means of educating physicians. Even informal actions are often serious (91). The propensity for inconsistency among such actions could be high, because informal actions are confidential. Such actions could be used selectively to avoid disciplining some physicians and not others.

Validity of the Indicator

About one-half of the formal disciplinary actions taken against physicians by State medical

Table 6-1.—Physician Disciplinary Actions by State Boards, 1986[a]

	License revocation	Probation	License suspension	Other regulatory action	Total
Alabama	4	4	0	7	15
Alaska	0	1	1	0	2
Arizona	6	4	2	55	67
Arizona[b]	0	6	2	16	24
Arkansas	1	1	4	11	17
California	34	69	18	43	164
California[b]	1	3	0	1	5
Colorado	4	10	10	8	32
Connecticut	5	2	2	8	17
District of Columbia	7	2	0	4	13
Delaware	2	0	0	0	2
Florida	22	23	30	117	192
Florida[b]	4	4	3	7	18
Georgia	16	42	15	35	108
Guam	0	0	0	0	0
Hawaii	6	1	1	2	10
Idaho	2	2	0	1	5
Illinois	10	25	38	47	120
Indiana	9	18	30	38	95
Iowa	16	7	3	9	35
Kansas	2	2	3	22	29
Kentucky	11	7	4	15	37
Louisiana	0	5	8	5	18
Maine	4	2	0	5	11
Maryland	1	5	5	15	26
Massachusetts	25	1	16	8	50
Michigan	3	0	11	2	16
Michigan[b]	0	0	5	2	7
Minnesota	3	22	5	11	41
Mississippi	2	3	3	16	24
Missouri	32	6	0	48	86
Montana	1	0	1	0	2
Nebraska	1	3	0	1	5
Nevada	2	1	1	3	7
Nevada[b]	0	0	0	0	0
New Hampshire	0	1	0	0	1
New Jersey	10	17	22	45	94
New Mexico	3	1	0	0	4
New Mexico[b]	0	0	0	0	0
New York	64	83	20	31	198
North Carolina	12	13	1	25	51
North Dakota	1	1	0	3	5
Ohio	24	20	14	51	109
Oklahoma	3	15	3	16	37
Oklahoma[b]	0	1	0	3	4
Oregon	2	19	7	20	48
Pennsylvania	9	7	11	34	61
Pennsylvania[b]	1	6	8	4	19
Puerto Rico	0	0	0	0	0
Rhode Island	2	0	1	4	7
South Carolina	6	5	5	10	26
South Dakota	1	0	0	8	9
Tennessee	6	3	1	11	21
Tennessee[b]	0	0	0	0	0
Texas	28	9	4	31	72
Utah	4	12	1	17	34
Vermont	0	0	1	6	7
Virgin Islands	0	0	0	0	0
Virginia State	15	18	6	51	90
Washington	0	8	3	22	33
Washington[b]	0	0	0	0	0
West Virginia	8	6	5	7	26
West Virginia[b]	0	0	0	0	0
Wisconsin	23	2	1	20	46
Wyoming	0	0	0	0	0
Total for year	458	528	335	981	2,302

[a] Except where designated, all boards take disciplinary actions against both allopathic physicians (M.D.s) and osteopathic physicians (O.D.s).

[b] This board takes disciplinary actions against osteopathic physicians (O.D.s) only.

SOURCE: B. Galusha and D.G. Breadon, "Official 1986 Federation Summary of Reported Disciplinary Actions," *Federation Bulletin* 75(2):41-46, 1988.

boards are on the grounds of inappropriate writing of prescriptions. Such infractions are the easiest to prove because of the exactness of prescription laws (658). Inappropriate prescribing and a physician's personal drug or alcohol abuse are the grounds for three-fourths or more of the disciplinary actions taken by State boards. Conviction for felony and fraud is among the most common of the remaining grounds for license discipline. A relatively small number of disciplinary actions are based on incompetence—the ground for discipline that would most clearly indicate poor quality of care.

If incompetence is strictly interpreted as the only violation that is a quality violation, disciplinary actions by State medical boards would not be a valid indicator of the quality of the medical care. A more liberal interpretation of incompetence to include inaccurate drug prescribing and drug and alcohol abuse is reasonable. The statistics just cited on types of violations present an incomplete picture of the importance of incompetence in disciplinary procedures. In addition, few medical practice acts identify incompetence as grounds for discipline, and the language of the acts that do is usually vague and difficult to interpret (694).[6] In addition, obtaining "clear and convincing evidence," of incompetence in most States is extremely difficult, time-consuming, and costly (239). Boards often use overprescribing of drugs and drug and alcohol abuse, which they have found often coincide with incompetence, as grounds for action instead of trying to prove incompetence (90,239,694,706,720). In particular, alcohol and drug abuse, characteristic of the impaired physician, and physical and mental illness can result in substandard performance and avoidable medical injury (636).

Several grounds for disciplinary actions are related to law and ethics. Many of these may not affect the technical aspects of quality but may influence interpersonal relations. The grounds vary greatly in seriousness and include conviction of a felony, conviction of a crime or felony related to medical practice, fraud in obtaining a license, violations of narcotics laws, violations of child abuse reporting acts, betrayal of professional secrets or privileged communications, and making untruthful or exaggerated claims relating to professional excellence or abilities (34,260). Other grounds for disciplinary action relate to charges of essentially economic violations, such as fraud regarding fees, fee-splitting, false or deceptive advertising, and overcharging or making false claims for reimbursement (34,260). Whether any, some, or all of these violations affect medical decisionmaking is not known, but to the extent that a violation affects an individual's trust in a physician's care, the ability of a physician to provide competent interpersonal care is compromised. People have different expectations of their physicians, and, depending on the type and seriousness of the violation, many people would not be comfortable going to a physician who had violated the law.

If one accepts that all violations that lead to formal disciplinary actions are quality violations, then such actions appear to possess validity as a measure of quality. The burden of proof for taking formal disciplinary actions rests with the State, and such actions usually must be based on "clear and convincing evidence," a difficult standard of proof. Due process safeguards are applied (70 Corpus Juris Sec. 43), and procedural aspects are sufficiently rigorous that the decisionmaking process is unlikely to be affected by external influences and the decisions are based on the evidence presented (260). The time taken to complete a formal disciplinary action—about 3 years—is indicative of the carefulness of the process.

Other factors operate in favor of protecting physicians' licenses. Inadequate funding and staff often limit States' ability to prepare their cases as well as the physicians' paid legal counsels.[7] Testimony from expert witnesses against the licensee has often been difficult to obtain because of a fear of civil liability for defamation (260,694).[8]

---

[6] No ground for discipline adequately describes the lack of professional ability or incompetence. The specific term varies among States and includes unprofessional conduct, gross incompetence, manifest incapacity, and malpractice and gross/repeated malpractice. All of these terms have no uniformly understood meaning.

[7] Andrew Watry, Executive Director of the Georgia State Board of Medical Examiners, reports that the Board's annual expenditures for legal fees for 60 actions is $80,000 to $100,000. A physician may spend as much as $50,000 to $100,000 in legal fees for one case (694).

[8] Professionals' concern might decrease as a result of the recent passage of the Health Care Quality Improvement Act of 1986 (Public Law 99-660). The act grants a limited immunity from damages un-

(continued on next page)

Thus, it is more than likely that physicians who have had formal disciplinary actions taken against them have violated State medical practice acts.

Nonetheless, the validity of formal disciplinary actions can be questioned, since the decisions of some boards have been overridden by the courts. Every State gives physicians the right to some type of judicial review of disciplinary actions taken against them to ensure that boards do not act in arbitrary, capricious ways or abuse discretion (260)(70 Corpus Juris Sec. 51). The courts have ruled against the boards in 30 percent of the cases brought before them (168,342)[9] on issues of constitutional rights, statutory interpretation, sufficiency of evidence, appropriateness of disciplinary action (260), and technical errors (169). Considering the number and range of reasons for overriding boards decisions, including technical errors, one can consider 30 percent a "fairly good record' (169).

Feasibility of Using the Indicator

Although information on formal disciplinary actions taken by State medical boards is available, consumers have limited access to it. Formal disciplinary actions are a matter of public record, and consumers can obtain information about actions taken against individual physicians by contacting State medical boards (190). Some boards even periodically report disciplinary actions to the news media (206), either directly or through newsletters, which almost a third of the boards now publish (206). Yet anecdotal information indicates that individuals and even representatives of health-related organizations are unaware of the availability of this information.

Another source of information on formal disciplinary actions by State boards, the Physician Disciplinary Data Bank operated by the Federation of State Medical Boards, is accessible only to organizations. The Federation's data bank includes information on formal disciplinary actions taken against physicians by its member State medical boards and other government authorities. The Federation of State Medical Boards sends monthly reports to its member boards and some private and public organizations on actions entered in the data bank during the preceding month (205). When the American Medical Association receives the Federation's monthly report, it informs all the State licensing boards under which a physician is licensed that the physician has been disciplined. The Federation also screens individual physicians' disciplinary histories upon request; in 1986 it answered 39,000 inquiries from member boards and other organizations (636). Organizations such as hospitals and insurance companies can contract with the Federation for information about disciplinary actions (90). Easier access to cross-State information will be available when the Federation completes a system for State medical boards to directly access the data bank (636).

(continued from previous page)
der Federal and State laws to individuals providing information to a professional review body regarding the competence or professional conduct of a physician unless they know the information is false.

[9] In a 1983 article, Derbyshire notes that the percentage was consistent for court decisions from 1902 to 1966 and from 1969 to 1979 (168). A similar percentage was found in an analysis of court decisions concerning actions taken against physicians who came before the Michigan Board of Medicine from 1977 to 1982 (342). More recent data are not available.

SANCTIONS RECOMMENDED BY PEER REVIEW ORGANIZATIONS AND IMPOSED BY HHS

In fulfilling its responsibility to assess and assure the quality of care provided to Medicare beneficiaries, HHS, upon recommendation of PROs, imposes sanctions on providers who fail to provide care that is medically necessary, appropriate, and of adequate quality.[10] The Secretary of Health and Human Services sanctions providers by imposing monetary penalties or exclusion from the Medicare program for specified periods of time.

The sanction process is initiated when a PRO physician finds that a quality problem exists and determines that a "substantial violation" or a

---

[10] See app. D for a comprehensive description of PROs.

"gross and flagrant violation" may have occurred.[11] A "substantial violation" is a pattern of care over a substantial number of cases that is inappropriate, unnecessary, does not meet recognized standards of care, or is not supported by the documentation of care required by the PRO. A "gross and flagrant violation" is a violation that has occurred in one or more instances and that presents an imminent danger to the health, safety, or well-being of a Medicare beneficiary, or unnecessarily places the beneficiary in a situation of high-risk, for example of substantial and permanent harm (638).

If a PRO believes that a provider's alleged violation was a "substantial violation," the PRO must give the provider two opportunities to discuss the allegations (see figure 6-1). Since the basic purpose of PROs is intended to be educational, the PRO first proposes corrective actions (e.g., requiring the physician to update skills by further education). If the quality problem is not corrected, the PRO recommends a sanction to the Office of the Inspector General of HHS. If the PRO believes that the provider's violation was a "gross and flagrant violation," the provider receives no opportunity to take corrective actions and only one opportunity for discussion before the PRO recommends a sanction (see figure 6-2).

In the case of "substantial violations" and "gross and flagrant violations," a provider is given 30 days notice and an additional opportunity to submit written comments before the PRO recommends sanctions to the Office of the Inspector General. The final decision about whether to sanction a physician is the responsibility of the Office of the Inspector General under authority delegated by HHS. The Office of the Inspector General decides if the medical evidence supports the decision of the PRO. If the decision of the Inspector General is to impose a sanction, a provider may appeal the decision to an HHS administrative law judge.

The intent of the discussion here is not to evaluate the effectiveness of the sanctioning process in identifying all providers of poor-quality care, but to evaluate whether PRO-recommended sanctions imposed by HHS are indicators of poor quality. As is true in the case of disciplinary actions taken by State medical boards, sanctions are expected to measure the overall performance of a provider. The hypothesized relationship, that PRO-recommended sanctions imposed by the Office of the Inspector General of HHS indicate providers of poor-quality care, has face validity. Since the Secretary of HHS is responsible for protecting the health and safety of Medicare beneficiaries, it is likely that beneficiaries and other consumers would consider physicians whom HHS fined or excluded from practicing in the Medicare and Medicaid programs[12] to be providers of poor-quality care.

Reliability of the Indicator

Sanctions result from actions taken by two different organizations, a PRO and the Office of the Inspector General of HHS. Because of variations in the process and criteria used to initiate sanctions among the 54 PRO programs, recommendations for sanctions by PROs on a national basis as an indicator of quality are not reliable (622). Furthermore, the criteria of "professionally recognized standards of care" that PRO reviewers use to assess the appropriateness and quality of providers' care are based upon typical patterns of practice within the PRO's geographic area or national criteria where appropriate (638). To the extent that PROs use local and regional standards of care in initiating sanctions, the criteria for assessing care can vary among areas. Since different criteria are likely to be used, the possibility of replicating sanction recommendations among PROs is low.

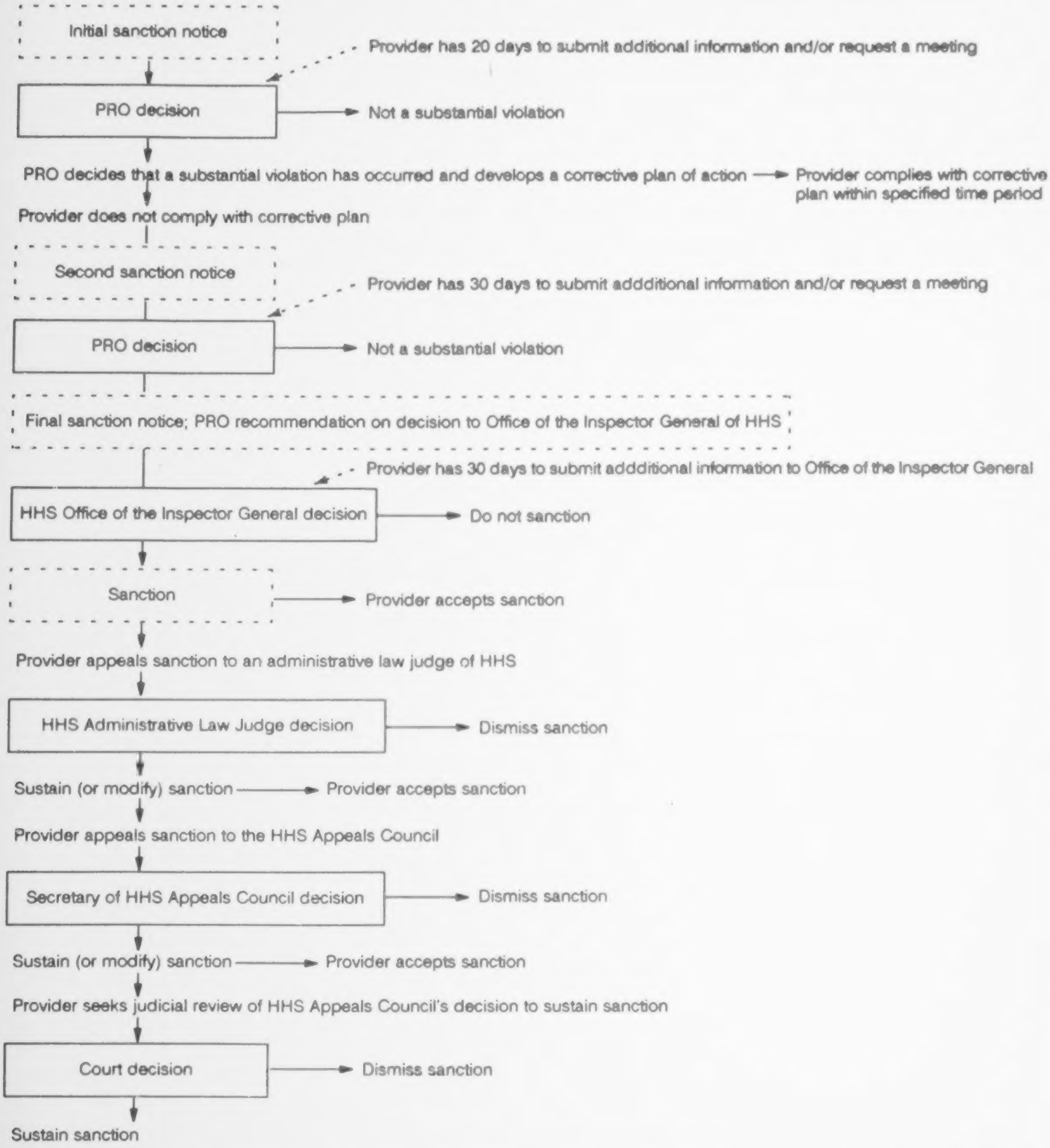
To the extent that a given PRO reviews similar cases in a similar manner, the PRO's recommendations for sanctions to the Office of the Inspector General may have a considerable degree of consistency as an indicator of quality. PRO rec-

---

[11] In addition to sanctions, PROs may also deny payment to providers. The Consolidated Omnibus Budget Reconciliation Act of 1985 (Public Law 99-272) gave PROs authority to deny payment for quality of care violations. As of February 1988, the final regulations on these denial notices had not been released.

[12] Public Law 100-93, the Medicare and Medicaid Patient and Program Protection Act of 1987, excludes physicians from Medicaid if they have been excluded from Medicare.

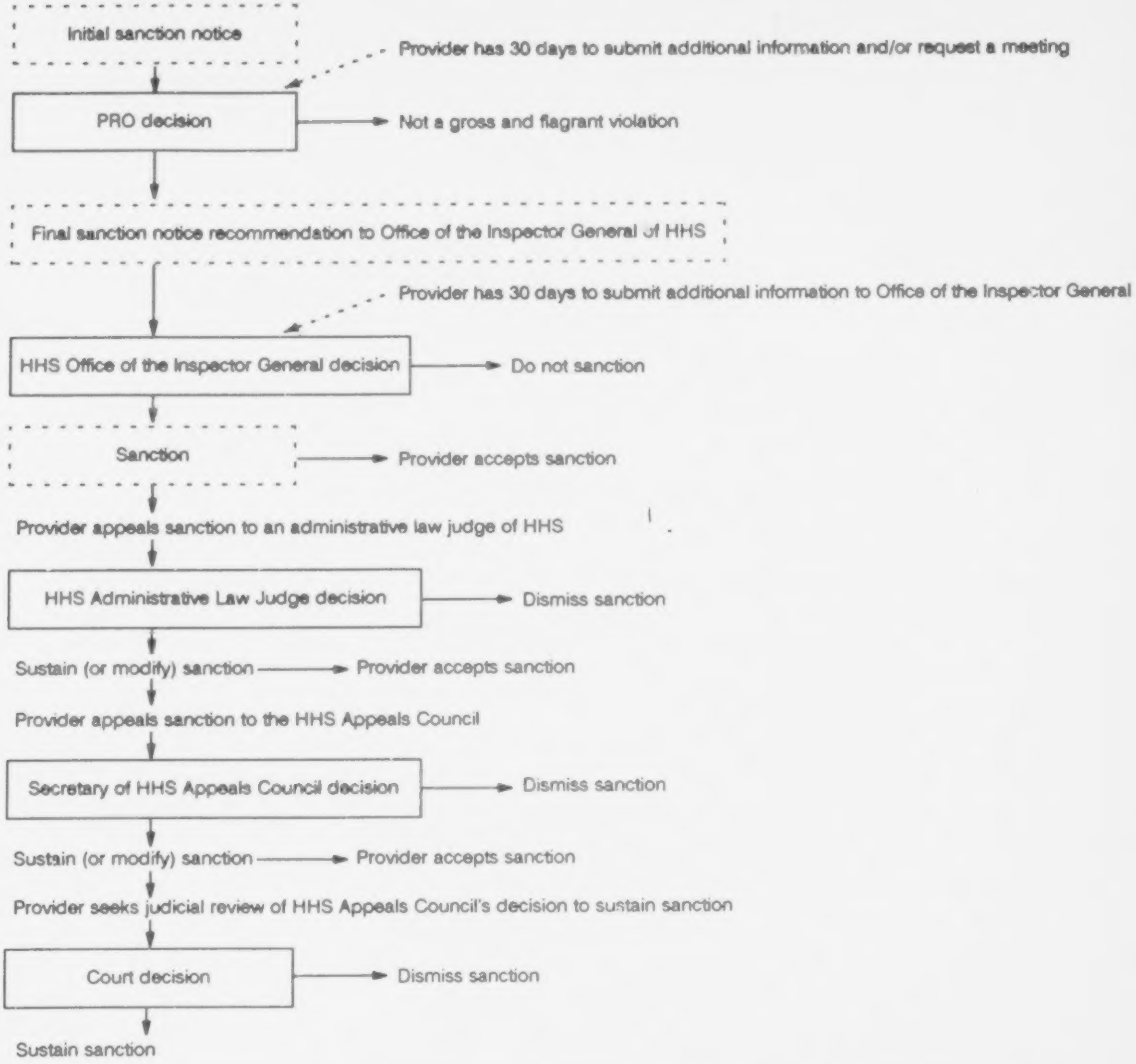
Figure 6-1.—Overview of the PRO/HHS Sanction Process for Substantial Violations[a]


Initial sanction notice
Provider has 20 days to submit additional information and/or request a meeting
PRO decision
Not a substantial violation
PRO decides that a substantial violation has occurred and develops a corrective plan of action
Provider complies with corrective plan within specified time period
Provider does not comply with corrective plan
Second sanction notice
Provider has 30 days to submit adddditional information and/or request a meeting
PRO decision
Not a substantial violation
Final sanction notice; PRO recommendation on decision to Office of the Inspector General of HHS
Provider has 30 days to submit addditional information to Office of the Inspector General
HHS Office of the Inspector General decision
Do not sanction
Sanction
Provider accepts sanction
Provider appeals sanction to an administrative law judge of HHS
HHS Administrative Law Judge decision
Dismiss sanction
Sustain (or modify) sanction
Provider accepts sanction
Provider appeals sanction to the HHS Appeals Council
Secretary of HHS Appeals Council decision
Dismiss sanction
Sustain (or modify) sanction
Provider accepts sanction
Provider seeks judicial review of HHS Appeals Council's decision to sustain sanction
Court decision
Dismiss sanction
Sustain sanction

[a] A substantial violation is a pattern of care over a substantial number of cases that is inappropriate, unnecessary, does not meet recognized patterns of care, or is not supported by the documentation of care required by the PRO

SOURCES: Adapted from U S Department of Health and Human Services, Health Care Financing Administration , "Peer Review Organization Manual," HCFA Pub No 19, Transmittal No 15, Baltimore, MD, May 1987; and R P Kusserow, Inspector General of the U S Department of Health and Human Services, testimony on the Peer Review Organization Process before the Subcommittee on Intergovernmental Relations and Human Resources, Committee on Government Operations, House of Representatives, U S Congress, Washington, DC, Oct 20, 1987

Figure 6-2.—Overview of the PRO/HHS Sanction Process for Gross and Flagrant Violations[a]


Initial sanction notice
Provider has 30 days to submit additional information and/or request a meeting
PRO decision
Not a gross and flagrant violation
Final sanction notice recommendation to Office of the Inspector General of HHS
Provider has 30 days to submit additional information to Office of the Inspector General
HHS Office of the Inspector General decision
Do not sanction
Sanction
Provider accepts sanction
Provider appeals sanction to an administrative law judge of HHS
HHS Administrative Law Judge decision
Dismiss sanction
Sustain (or modify) sanction
Provider accepts sanction
Provider appeals sanction to the HHS Appeals Council
Secretary of HHS Appeals Council decision
Dismiss sanction
Sustain (or modify) sanction
Provider accepts sanction
Provider seeks judicial review of HHS Appeals Council's decision to sustain sanction
Court decision
Dismiss sanction
Sustain sanction

[a] A gross and flagrant substantial violation is a violation that has occurred in one or more instances and presents an imminent danger to the health, safety, or well-being of a Medicare beneficiary

SOURCES: Adapted from U.S. Department of Health and Human Services, Health Care Financing Administration, "Peer Review Organization Manual," HCFA Pub. No. 19, Transmittal No. 15, Baltimore, MD, May 1987, and R.P. Kusserow, Inspector General of the U.S. Department of Health and Human Services, testimony on the Peer Review Organization Process before the Subcommittee on Intergovernmental Relations and Human Resources, Committee on Government Operations, House of Representatives, U.S. Congress, Washington, DC, Oct. 20, 1987 ommendations for sanctions go through a number of reviews before they are sent to the Office of the Inspector General. The first round of physician review offers chances for great inconsistency. Similar cases could be reviewed by different physicians who for the most part use implicit criteria in deciding to initiate a sanction. Furthermore, some PROs have expressed concern that inadequate funding makes them unable to recruit, train, and retain qualified physician reviewers (491).

Nonetheless, subsequent reviews can increase the chances that a recommendation for a sanction for a similar violation is replicable within a PRO. The number of additional reviews varies among the PROs. In Iowa, for example, before a sanction is recommended to the Office of the Inspector General, the case is reviewed by a 15-member quality assurance committee; a 15-member comprehensive review committee; and the board of directors of the PRO, composed of 29 physicians, a business representative, a dentist, a nursing home owner, an administrator of a small hospital, and an administrator of a large hospital (405). Before a sanction recommendation is made to the Office of the Inspector General, the California PRO involves a regional medical director, the associate medical director, the medical director, the monitoring committee, the chief executive officer, and the board of directors (435). In all PROs, final review by the PRO's board of directors is required before a formal recommendation is made to the Office of the Inspector General.

To the extent that a PRO's board of directors is stable in membership, that a consensus process is used in arriving at decisions, and that members are consistent in their rulings, reliability is increased. If precise guidelines were used by boards of directors in arriving at recommendations for sanctions, the replicability of their decisions could be increased. More exact guidelines were provided in May 1987 as the result of an agreement among the American Association of Retired Persons, the American Medical Association, the Office of the Inspector General, and the Health Care Financing Administration (HCFA) to specify and standardize the procedures PROs use in recommending sanctions (164).[13]

Since the imposition of sanctions is, for the most part, a function of the Office of the Inspector General of HHS, the additional reviews the Office conducts before a provider is sanctioned are crucial in establishing the reliability of sanctions. Federal regulations are specific about what steps the Office should take in arriving at a sanction decision, but do not describe how the steps should be executed (42 CFR 1004.90 [1986]). The same small number of Office personnel, representing the medical and legal professions, are involved in considering whether a provider has violated his/her obligations and in determining an appropriate sanction, and a single individual within the Office of the Inspector General is responsible for the final determination to sanction a provider (375).

Validity of the Indicator

It is not clear whether all sanctions are initiated on the basis of quality-related problems.[14] Recommendations for sanctions are initiated by PROs when a provider's services: 1) are not provided economically and are not medically necessary, 2) are not of a quality that meets professionally recognized standards of health care, and 3) are not properly documented (638). Although provision of unnecessary services could be classified as a quality issue, insufficient documentation is most likely due to inadequate recordkeeping, which may or may not be associated with poor-quality care.

---

[13] The recommended procedures include specifying model letters that PROs will send to physicians and hospitals during the sanction process; ensuring that no physician member of a PRO making a final sanction determination against a physician has a bias against or is in competition with the subject physician; permitting an attorney to accompany a physician to certain meetings required during the process; permitting the attorney to make opening and closing remarks and to assist the physician in presenting the testimony of expert witnesses who may appear on the physician's behalf; making a verbatim record of such meetings with a copy made available to the physician: and permitting the physician to submit additional relevant information to the PRO within 5 working days after the meeting (164).

[14] The Health Care Financing Administration collects data on the number of sanctions initiated by PROs because of potential "substantial violations" and "gross and flagrant violations," but does not have information on the grounds for the initiation of sanctions (228). The Office of the Inspector General does have the information but does not generally distribute it. The Office of the Inspector General has provided the information to at least one consumer advocacy group.

Of more importance is the fact that, if unnecessary and inappropriate services and premature discharges are perceived as quality concerns, almost all of the sanctions imposed by the Office of the Inspector General upon recommendation of PROs have been for quality violations. In fact, 78 of the 79 sanctions the Office imposed by September 1987, were on quality-based grounds. One sanction was based exclusively on grounds of improper documentation (375).

Furthermore, the possibility of sanctioning physicians who do not provide poor care is slight, an observation that suggests that PRO-recommended sanctions imposed by HHS are valid indicators of quality. An extensive weeding-out process takes places before PRO-recommended sanctions are imposed by HHS, and only a few sanctions have been imposed. From the 30 million hospital discharges involving Medicare beneficiaries from the beginning of October 1983 to the end of December 1986, PROs identified 6,500 discharges involving 2,500 providers as having potential quality-of-care or utilization problems (360). The great majority—over 97 percent—of the cases were resolved at the PRO level by PROs working with providers during the steps of the process and were not referred to the Office of the Inspector General. Most deficiencies were corrected by educational or corrective actions, and through December 1987, only 151 cases were referred to the Office for review and final action. Not all of the 151 cases that were referred were held to be sustained in law or by medical evidence. Only 61 resulted in exclusion from the Medicare program (60 physicians and 1 hospital); 26 cases resulted in a monetary penalty; 8 cases are now under review; and 2 physicians have died (661). Many of the sanctions that the Office rejected were rejected because of procedural issues (e.g., the PROs were late in submitting documentation or the documentation was not complete) (360).

Physicians who are sanctioned are often cited for multiple violations. One physician, for example, was sanctioned on the basis of 22 cases of deficient care (713). Indeed, the 11 physicians who were sanctioned by exclusion from Medicare in the period February 8 to July 2, 1987, were responsible for "gross and flagrant violations" in the care of 48 patients. Physicians who are fined are also likely to have committed one or two serious violations (501).

Another way to determine if physicians who provide standard care are being safeguarded from sanctions is to examine if they are given appropriate due process. The sanction process attempts to balance the interest of HHS in protecting the health and safety of Medicare beneficiaries with the due process rights of providers. As a peer review process, the system does not have the extensive safeguards characteristic of the judicial process. Nonetheless, physicians have the opportunity of submitting information, being heard before two administrative bodies (the PRO and the Office of the Inspector General) before a sanction is imposed, and of appealing the imposition of a sanction.

As noted earlier, regulations require that PROs allow physicians to submit information and meet *once* with the PRO if they are alleged to have committed a gross and flagrant violation(s) and *twice* if the violation is a substantial violation. After the PRO recommends a sanction, the Office of the Inspector General conducts an independent review of the PRO report and any additional information submitted by the physician under consideration. If the Office agrees with the PRO and also finds the physician unwilling or unable to comply with statutory obligations,[15] it will sanction the physician, [16] either by excluding the physician from the Medicare program or by imposing a monetary penalty. The Office's decision can be appealed, in which case a hearing is held before an administrative law judge of HHS. This hearing is the first time in the process that a full evidentiary hearing is held (360). This decision may then be reviewed upon request by the HHS Appeals Council. If dissatisfied with the result, physicians have the right to seek further review of their cases in the court system (see figures 6-1 and 6-2).

---

[15] The "unwilling and unable" condition has been questioned as being ambiguous and an impediment to protecting patients from providers of substandard care (360).

[16] The Office of the Inspector General of HHS has 120 days to accept or reject the recommendation, or a sanction is imposed. To date, the Office of the Inspector General has always acted on recommendations within the allotted time (375).

A few cases have gone to district courts, and some of them have been appealed. The appeals courts have upheld the adequacy of due process in the PRO sanction process (125,276,674). In its ruling, the 4th Circuit Court noted that the PRO-initiated sanctions process affords providers appropriate due process, since the Government's need to protect Medicare beneficiaries from poor-quality care is compelling. Disagreement with the adequacy of the process continues, however, in both the medical and legislative communities. Some accommodation was made in the Omnibus Budget Reconciliation Act of 1987 (Public Law 100-203), which allows physicians in certain underserved areas to continue to practice, although sanctioned, during the administrative review process.[17]

Feasibility of Using the Indicator

Although sufficient data exist for purposes of formulating PRO-recommended sanctions imposed by HHS as an indicator of the quality of care, many consumers may find it difficult to gain access to the information. If the Office of the Inspector General imposes a sanction, a notice is published in a newspaper in the PRO area advising the public of the Government's action.[18] The May 1987 agreement between the Office of the Inspector General, the American Medical Association, the American Association of Retired Persons, and HCFA, discussed earlier, included a stipulation, subject to regulatory change, that physicians and other providers are to notify their Medicare patients that they have been sanctioned in lieu of newspaper publication of this fact. As of February 1988, the regulations were under development.

Anecdotal information indicates that the current publishing requirement has not been implemented in a way that provides easy access for consumers to information about sanctioned physicians. It is said, for example, that newspapers with small circulations are often used, and notices are placed in small type, often in the public notice section. There are potential problems with the new method as well. It is not clear that having physicians privately inform their current or potential Medicare patients that they have been sanctioned will increase the effectiveness of providing consumers with information on sanctions. On the one hand, Medicare beneficiaries would not have to seek out the information. On the other hand, sanctioned physicians will have conflicting interests between defending their practices and their legal obligation to provide information. Thus, Medicare beneficiaries may not receive as complete an explanation about the grounds for a sanction from the sanctioned physician as from a newspaper notice. In addition, automatic notification concerning sanctioned physicians will not be available to non-Medicare consumers (501).

---

[17] The Omnibus Budget Reconciliation Act of 1987 (Public Law 100-203) made some changes in the review process. The act provides that in rural health areas with health manpower shortages and in counties with fewer than 70,000 people, physicians seeking to overturn a decision that excludes them from Medicare for failure to furnish medical care of acceptable quality may continue to practice during the administrative review process, unless an administrative judge determines that continued practice would pose a serious risk to Medicare beneficiaries.

[18] Regulations require that the Office of the Inspector General "notify the public by publishing in a newspaper of general circulation in the PRO area a notice that identifies the sanctioned practitioner or other person, the obligation that has been violated, the sanction imposed and if the sanction is exclusion, the effective date and duration" (42 CFR Sh. V (10)-1986 ed.).

MALPRACTICE COMPENSATION

A malpractice suit indicates that a patient is dissatisfied with care received from a provider. There is some evidence that patients who are satisfied with interpersonal aspects of care are less likely to sue their physicians than patients who are dissatisfied with these aspects.[19] The analysis in this chapter assumes that a malpractice suit suggests that the physician has some deficiency in interpersonal aspects of care. The question explored here is whether malpractice compensation is a reasonable indicator of poor technical performance.

Patients usually use the tort liability system to obtain monetary compensation for medical injury. The process of determining liability is initiated when a patient or provider identifies a medical, possibly negligently induced, injury. Sometimes a "warning" file may be opened by an insurer on the basis of a report from an insured provider. The next step is likely to be a claim received directly from an injured claimant or claimant's representative, which may accompany or soon be followed by a lawsuit. The lawsuit may or may not be resolved in favor of the patient. If the lawsuit is resolved in favor of the patient, the patient may receive medical compensation either through a jury verdict or through negotiated settlement with the physician's insurance company, usually prior to an actual courtroom proceeding. Even after a jury verdict, the trial judge may alter or overturn the verdict, and appeals may be made. Many awards are reduced before actual payment is made (159).

Only some medical injuries, or adverse medical outcomes, that occur are the result of providers' failing to conform, through omission or commission, to current standards of medical care (449,636). Other adverse medical outcomes are unavoidable results of insufficient medical knowledge about the natural course of some conditions and unexpected effects of diagnostic and therapeutic procedures (449,636).

Few studies have attempted to determine the occurrence of medical injuries and fewer still the proportion that are possibly negligently induced. A pilot study in 1973 found that medical injuries occurred in nearly 8 percent of the cases reviewed and estimated that medical injuries due to medical negligence occurred in 2.3 percent of cases reviewed (494). A study of about 21,000 hospital records of California hospitals performed in the mid-1970s concluded that 1 out of every 20 admissions (5 percent) resulted in an injury caused by medical treatment (114). Seventeen percent of the injuries caused by medical treatment, or 1 out of every 126 hospital admissions (0.8 percent), were estimated to be caused by legally provable negligence. A more recent analysis found that almost 1 percent of hospital admissions are associated with poor care that results in temporary or permanent disability or death (159).

The discussion here will focus on the reliability and validity of court awards as an indicator of the quality of care, in part because from a consumer's perspective such awards would have face validity. In order for medical malpractice to be established in court, one must prove the existence of a duty of the physician to the patient, the existence of an applicable standard of care, negligence or the failure of the provider to meet the standard of care, injury or damage to the patient, and the determination that the proximate cause of injury to the patient was the physician's failure to meet the standard of care (636). Most people would consider a physician who has been found liable of malpractice in a court action, particularly for a number of cases over a period of time, to be a provider of poor-quality care.

---

[19] Obstetricians/gynecologists and medical specialists who report spending more time with their patients per office visit than similar physicians, on average, incur fewer claims than physicians who spend less time (6).

The other major form of malpractice compensation is negotiated settlements, that is, payment without a judicial determination. Malpractice settlements are often made for reasons unrelated to quality that are usually unknown to the general public.[20] The lack of information does not mean that any claim for which settlement was negotiated is not without meaning as an indicator of quality. Indeed, it can be argued that cases with large settlements are settled out of court because negligence can be proven. It is likely that cases involving small settlements of $20,000 to $40,000 may not go to trial for reasons of efficiency as well as reasons of negligence. Furthermore, negotiated settlements are the more important form of claims settlement, since 90 percent of medical malpractice claims are settled before trial, and of those settled with payment approximately 97 percent are closed as a result of a negotiated settlement (625). The reliability and validity of negotiated settlements as indicators of the quality of care cannot be evaluated, however, because the negotiation process is confidential.[21]

Reliability of the Indicator

The many variations across the country in the tort system governing malpractice cases—including variations in laws, judges, and juries—make it unlikely that court awards are reliable as an indicator of quality on a national level. There are fewer variations within States, because medical malpractice claims are resolved through State court systems and under State statutes. Even within States, however, many judges and juries are involved in malpractice cases, and not all judges place the same interpretation upon the law.

Within a judge's courtroom, a judge's awards may be consistent. In addition, indirect evidence suggests that jury awards might have some reliability as an indicator of the quality of care. Although the consistency of verdicts among juries has not been studied, the consistency in verdicts between judge and juries has been examined in both criminal and civil cases. In 3,576 criminal trials and about 4,000 civil trials, both judge and jury agreed on the verdict 78 percent of the time (338). These findings might have positive implications for consistency among juries. In general, studies find that the rate of agreement among participants in complex human judgments, such as scientific peer review panels and decisions of practicing physicians and judges, ranges from 55 to 80 percent (see table 6-2). Nonetheless, the limited boundary of one judge's courtroom, within which jury awards might have some consistency, works against the usefulness of jury awards as an indicator of the quality of care.

Validity of the Indicator

Individual Awards for Malpractice

Court awards for malpractice as a measure of the quality of care would appear to have some validity as indicators of quality in that compensation is supposed to be awarded for negligence only. Other concerns, such as fraud and abuse, are not at issue. In addition, a judgment in favor of a patient, in theory, means that a physician's negligence has been proven. Nonetheless, in weighing the evidence it appears that—except in extreme cases, such as amputating the wrong limb—individual jury awards are not indicators of a physician's performance.

On the one hand, the difficulty and length of time involved in filing and resolving malpractice claims, the formal process of the litigation, and the small number of cases that are resolved in favor of the patient/claimant appear to support the contention that physicians who have been found liable of malpractice are providing poor-quality care. Although these features have to do with

---

[20] Settlements are usually agreed to in cases whose resolution is clear (88). They are often made for reasons other than physician negligence, including court congestion, variation in the interest of liability insurance carriers in settlements, probability of success in a particular court before specific judges, credibility of both plaintiffs and defendants as witnesses (461), the cost of protracted litigation compared to early settlement, the ability of the plaintiff's counsel, the sympathy aroused by the plaintiff, aggravated fact situations that would inflate the award, the amount of the awards for a similar injury in the jurisdiction, the personalities of the key witnesses, the desire to avoid publicity of a trial, and the existence of a statutory requirement to submit the claim to a pretrial panel (636). Indeed, some experts contend that settlements are not directly related to a finding of malpractice, i.e., negligently induced medical injury (636).

[21] The section does not consider malpractice claims that have not been resolved, because such claims represent an accusation of wrongdoing with no knowledge of the truth or falsity of the claim.

Table 6-2.—Interjudge Consistency in Complex Human Judgments

Decisionmakers	Stimulus	Decision	Rate of agreement between 2 judges (%)
National Science Foundation vs. National Academy of Sciences peer reviewers	150 Grant proposals submitted to the National Science Foundation	To fund or not to fund (half funded by the National Science Foundation)	75
7 Employment interviewers	10 Job applicants	Ranked in top 5 or in bottom 5	70
4 Experienced psychiatrists	153 Patients interviewed twice, once by each of 2 psychiatrists	Psychosis, neurosis, character disorder	70
21-23 Practicing physicians	3 Patients-actors with presenting symptoms	Diagnosis: correct or incorrect	66, 77, 70
		Probability of agreement (both correct or both incorrect)[a]	55,65,57
3,576 Judge-jury pairs	3,576 Jury trials	Guilty or not guilty	78
12 Federal judges	460 Presentence reports (at sentencing council)	Custody or no custody	80
8 Federal judges	439 Presentence reports (at sentencing council)	Custody or no custody	79

[a] Inflated because physicians could also be inaccurate in different ways.

SOURCE: S.S. Diamond, "Order in the Court: Consistency in Criminal Court Decisions," *The Master Lecture Series Vol. II: Psychology and the Law*, C.J. Scheirer and B.L. Hammonds (eds.) (Washington, DC: American Psychological Association, 1983). Copyright 1983 by the American Psychological Association. Adapted by permission of the publisher and the author.

adherence to procedural requirements rather than with the substance of claims, one could argue that they diminish the possibility that physicians who are found liable will not in fact have been negligent. Of course, some of the physicians who are not found liable of malpractice may in fact have been negligent.

The fact that very few injured people bring a malpractice claim (87) illustrates the difficulty of the process. A recent pilot study of the prevalence of public perceptions of medically induced illness in Maine concluded that of the 42 respondents that had reported that they or a close relative had a medically induced injury, 2 discussed the incident with an attorney and only 1 initiated a suit (430). A more comprehensive analysis estimated that claims are filed for only a small percentage of negligently induced injuries. Extrapolating from a 1977 California Medical Association/California Hospital Association study and 1974-76 data collected by the National Association of Insurance Commissioners, the researcher estimated that about 1 malpractice claim was filed for every 10 potentially valid claims (159).

An attorney has to be convinced of the merits of a case to take the case, because most attorneys in malpractice litigation cases are paid only if their client wins (i.e., they work on a contingency fee basis). Since attorneys generally receive a percentage of the award, most are concerned with potentially successful claims that are likely to result in a substantial award. Although it is obvious that the number varies among lawyers, a dated survey found that a claimant has less than one chance in eight of convincing an attorney to take a medical malpractice case (449).[22]

The extensive time required is another illustration of the rigor of the claims resolution process. The median length of time from claim filing to complete disposition against all the providers involved is 19 months: the median time for paid claims is 23 months. In general, the more severe and the more costly cases take a longer period to resolve (623).

Furthermore, during litigation, the substantive and due process rights of participants are protected. Formal rules of evidence control the admission of unreliable or prejudicial testimony, and compensation depends upon proving the provider at fault (449). Standards of care are generally in favor of the defendant (87). Providers are judged by peer standards, and juries are instructed to assess and choose among the medical opinions presented and not impose their own opinion of the care. Finally, only a small percentage of claims filed are closed with a court award. A study of 73,500 claims closed in 1984 found that 24,630 (43.7 percent) were closed with payment; of the 24,630, only 608 (2.5 percent) were closed with a court verdict either before or after appeal (622).

---

[22] Newer quantitative data are not available.

On the other hand, in reality, numerous other factors not related to the quality of medical care influence jury awards. Such factors include the effectiveness of the attorneys (611); the ability of the jury and expert witnesses to assess medical responsibility (611,636); the effect of race, sex, and perceived economic status on the jury (486); the effect of the passage of time from incident to verdict on the quality of the evidence (317); and the selective recall of witnesses (486); the effect of the extent of the injury and its obviousness (e.g., when surgical instruments are left within a body) (159); and the effect of the number of defendants (the chance of a physician's receiving an adverse judgment approximately doubles when a case involves multiple defendants) (159). It is not known whether some of these factors lead to increased or decreased accuracy in the outcomes of medical malpractice litigation.

In addition, individual jury awards are inaccurate indicators of specific physicians who provide substandard care, because multiple physicians may be defendants in any one case. Physicians who have had only peripheral involvement with a supposed negligently induced injury may be involved in the jury award. Heads of departments, for example, are often held legally responsible for the actions of the residents in their department, even though they were not present at the time of an incident; the same may be true of residents who played only a small part in a complex procedure.

Another challenge to the validity of malpractice compensation as an indicator of the quality of medical care is that malpractice litigation depends to a large extent on the lack of criteria regarding poor-quality care. The disagreements about what constitutes real malpractice are longstanding and serious and need extensive research before resolution.

Physician Profiles

A successful malpractice suit might indicate that a physician made an inadvertent error that had serious consequences for the patient or it might be one instance of a dangerous practice pattern of a physician that poses a risk in future patient encounters. There is a lack of empirical evidence to indicate which applies. Some would argue that findings of negligence in a number of malpractice cases indicate that a physician is delivering substandard care. Although this argument may seem intuitively correct, evidence to disprove it is also lacking.

A report of an analysis of Maryland data from 1960 to 1970 noted that a physician's being sued more than once could be attributed as much to chance as to poor practice, but the authors warn against generalizing the data to the entire country (101). A hypothetical informal statistical analysis confirmed the above finding (443). The analysis assumed that all physicians were similar and all patients were similar and that all cases were independent of each other. Yet in practice, physicians practice in different specialties and even, within a specialty, see different types of cases and different numbers of patients. Physicians who are frequently sued may be technically excellent but may be treating difficult cases and using high-risk procedures. In the absence of knowledge about patient and practice characteristics, the relationship of a physician's quality of care to multiple malpractice suits cannot be determined.

It is clear, however, that liability experience is not random with respect to specialty, and that some specialties have more malpractice claims than others. The specialists most often named in malpractice actions are obstetricians/gynecologists, general surgeons, and orthopedic surgeons; the percentage of claims paid is highest for pathologists, urologists, otolaryngologists, and obstetricians/gynecologists (623). These specialties employ invasive technologies with greater chances of doing serious harm. The many suits against obstetricians may also reflect heightened expectations on the part of consumers about what can be done with procedures such as fetal monitoring or amniocentesis rather than anything to do with the technical aspects of quality.

Studies also show that fairly few physicians account for a large share of medical malpractice claims. One study reported that 1 percent of physicians were responsible for 25 percent of paid claims and 20 percent of physicians had three or more paid claims in 10 years (301). Another study found that about 42 percent of physicians with claims in one year had previous claims against them (623). Since such percentages reflect the differences in malpractice experience among specialties, they do not necessarily mean that these physicians are providing substandard care.

Certain physicians in certain specialties have more claims than expected by chance (301,529, 675). In looking at large claims, researchers found that in some specialties, some physicians did not have more claims than expected (675). In other specialties, including internal medicine and anesthesiology, some physicians had disproportionately more claims than others; however, the difference could be accounted for by differences in practice level. This finding indicates that the past experience of individual physicians in certain specialties may be a valid measure of the individual's exposure to claims in the future and may be used to set malpractice premium rates. The lack of information about the characteristics and numbers, however, of the patients and cases seen by physicians compromises the ability to use medical malpractice experience as a valid indicator of substandard care provided by individual physicians.

Currently, the frequency and severity[23] of claims against individual practitioners are taken into account in quality-of-care evaluations carried out by certain hospitals for the purposes of peer review and by certain State licensure and credentialing organizations. The impetus for this new practice can be traced to lawsuits in Arizona and California, where hospitals had been held responsible to patients when lawsuit information of attending physicians was not considered when medical staff committees determined whether to grant hospital privileges (197,504). The frequency and severity of claims are also used by certain insurance organizations to evaluate physicians who are applying for malpractice insurance coverage or renewal and to identify physicians for risk management and quality assurance review and remediation (30).

Feasibility of Using the Indicator

The remarkable limitations of available data on malpractice litigation contravene the feasibility of using medical malpractice compensation as an indicator of the quality of care. The major source of data on settlements and jury awards is claims closed by insurers writing malpractice insurance,[24] and data from this source are expensive to collect and limited in usefulness. One reason that the usefulness of the data is limited is that insurers do not have a standard definition of claims and count claims differently.

Systematically collected data on the number of paid malpractice physician claims are not readily available; also not readily available are data on the frequency of malpractice claims that involve multiple providers and the identity of the defendants in multiple-defendant malpractice suits. Health insurance data that link procedures performed to individual physicians would be helpful in addressing the issue of the relationship of multiple settlements to extent of practice. In most instances, such information is not available. Data that identify physician performance that results in negligent actions and malpractice claims are not available. Without such information, it is not possible to relate malpractice compensation to negligence. To obtain such information, costly medical record reviews would be needed in addition to malpractice claims information.

Incomplete information on medical malpractice judgments is compiled at present, but even this information is not readily accessible to consumers. A malpractice judgment is a final court decision, and like any other court record, it is public. Some State laws require reporting of malpractice judgments to medical licensing boards. If the State has a Freedom of Information Act, the information is available through a Freedom of Information request (578). Although an ongoing source of data on jury verdicts is the privately published Jury Verdict Reporter Newsletters, which cover many metropolitan areas, such publications are expensive, and it is unlikely that individual consumers subscribe to them. Information on out-of-court settlements is not publicly available.

---

[23] Severity is related to frequency. Potentially high damages are more likely to prompt a claim than are low ones, and anyone specializing in high-damage cases, such as obstetricians/gynecologists, is likely to generate higher frequency claims than other specialists.

[24] The last study of national claims identified a universe of 102 malpractice insurers in the United States in 1983 (623).

CONCLUSIONS AND POLICY IMPLICATIONS

The causal relationship between license discipline, sanctions imposed by HHS upon recommendations by PROs, and malpractice compensation on the one hand and quality of care on the other has not been the subject of scientific examination. Since the interpretation of such relationships relies on deductive reasoning from descriptive information, findings are not firmly conclusive. Nonetheless, tentative conclusions can be made and directions for policy and research offered.

Disciplinary Actions by State Medical Boards

Of the three potential indicators examined in this chapter, formal disciplinary actions taken by State medical boards can currently be used with the greatest degree of confidence in identifying physicians who provide substandard care. Although the reliability of disciplinary actions is not clear, the deliberateness of the disciplinary process and the safeguards of physicians' rights to legal due process appear to ensure that the actions indicate infractions of State medical practice acts. Some people do not consider all infractions of State medical practice to be quality problems, however, because the scope of medical practice acts is broad and infractions of the acts include inaccurate drug prescribing, substance abuse, and criminal actions as well as incompetence. For those consumers who believe that quality in providing medical care is affected by a physician's character and not confined to the physician's technical skills, formal disciplinary actions taken by a State medical board would be a fairly good indicator of poor-quality care. For those consumers who limit their assessment of the quality of medical care to how physicians provide medical care, formal disciplinary actions generally would be an inexact indicator of poor-quality care. For all consumers, formal disciplinary actions that are taken on grounds of incompetence are adequate, albeit not perfect, indicators of substandard care.

If the reliability of formal disciplinary actions were better established, individuals and organizations could use this indicator with greater confidence. In order to increase reliability, an essential step would be to open up to public examination the processes that State medical boards use in disciplining physicians. Public scrutiny would also permit a better understanding of informal disciplinary actions and exactly when, why, and how they are taken and enforced. Their relationship to formal disciplinary actions and to poor care has not been examined. The validity of disciplinary actions as a quality indicator could be improved if all State medical practice acts included incompetence as a ground for disciplinary action, precisely defined the meaning of the term, and supplied guidelines for the actions applicable to the violation.

Although consumers can obtain information about formal disciplinary actions taken against individual physicians by contacting State medical boards, most consumers do not know this. Information would reach more consumers if more State boards would publicize their actions widely, and if State boards that currently supply information would increase their dissemination activities. Without additional funding, most State medical boards would have difficulty assuming the additional costs associated with providing information to the public. Most of the boards are under extreme financial constraints due to increasing investigatory and disciplinary activities (361). If dissemination of such information is a desirable government responsibility, additional State funding is needed. Federal funding is another possibility, although many concerned individuals believe that it would interfere with States' prerogative to license physicians (190).

Another source of information on formal disciplinary actions taken by State medical boards is the Physician Disciplinary Data Bank operated by the Federation of State Medical Boards. Reporting of disciplinary actions by State medical boards to the Federation is voluntary, but all States participate in the Federation's data bank. Through monthly reports and through direct access to the data bank, the information is disseminated to State medical boards and other organizations; it is not disseminated to individuals. Some would argue that the usefulness of individual access to the information in the Federation's data bank is questionable. Although organizations such as third-party payers require updates on disciplinary actions taken against many physicians, most individuals are interested in information concerning one or more physicians at one point in time, and that information can be obtained from State medical boards. The Federation charges for its services, and the charges might be high for most people. In addition, the Federation does not verify the accuracy of the information that the States report. Organizations are expected to use the information in the Federation's data bank as a starting point for more intensive inquiry—a course which many individuals might not be willing or able to pursue.

The national data bank mandated by the Health Care Quality Improvement Act of 1986 (Public Law 99-660) is a potential source of information on disciplinary actions.[25] State medical boards are to report disciplinary actions to the data bank, but are not mandated to actively obtain information concerning other boards' disciplinary actions. It appears the data bank will include the same license discipline information now available in the Federation's Physician Disciplinary Data Bank, but will add new information on malpractice compensation and adverse actions taken by hospitals regarding physicians' privileges. National confidentiality requirements will not override State legislative requirements of confidentiality (706).

Sanctions Recommended by PROs and Imposed by HHS

It is likely that sanctions imposed by the Office of the Inspector General of HHS on the recommendation of PROs are indicators of substandard quality of care. Available evidence about the sanctioning process suggests that recommendations for sanctions are consistent within a PRO area and are imposed consistently by the Office of the Inspector General. Such sanctions are valid indicators of physicians and hospitals that provide unnecessary services and substandard care. But evidence is very scanty and the sanctioning process is new and evolving. Although consumers could use such sanctions as an indicator of poor-quality care at this time, the indicator needs continuous evaluation.

The reliability and the validity of sanctions as an indicator of quality could be assessed with greater accuracy if information about the processes used by PROs and the Office of the Inspector General were available. It is clear that there is great variation in the approaches used by PROs in assessing quality, the number of groups within a PRO that review a case, and the number and types of intervention steps and amount of time between the identification of a quality problem and sanctioning (623,661). Yet little is known about how individual PROs make sanction recommendations and how the Office of the Inspector General executes the steps in arriving at a sanction decision. It would appear that the use of precise guidelines by the boards of directors of PROs in recommending sanctions to the Office of the Inspector General and the standardization of professional guidelines of care would allow consumers to rely more heavily upon PRO/HHS sanctions as a quality indicator.

The potential usefulness of this indicator of the quality of care suggests that a policy requiring oversight of the effectiveness of actions to disseminate information on sanctions is warranted. A new method has been agreed upon, and once regulations are promulgated, providers will have to notify their Medicare patients of sanctions. Sanctioned physicians may be hesitant about providing complete information to their Medicare patients, and their non-Medicare patients may not be informed at all. Although private publications, specifically the newsletter published by the Public Citizen Health Research Group, periodically publish the names of sanctioned physicians and analyze the grounds for sanctions, these publications do not reach all Federal beneficiaries.

---

[25] The national data bank did not receive funding for fiscal year 1988, although it is in the President's budget for fiscal year 1989.

A serious gap in availability of information is the lack of a central source for obtaining information about physicians who have been sanctioned by HHS as a result of PRO recommendations. As mandated by the Health Care Quality Improvement Act of 1986 (Public Law 99-660), the proposed national data bank is not intended to include information on sanctions imposed by HHS that result from PRO recommendations. In any event, the information in the data bank will not be publicly available.

Malpractice Compensation

Medical malpractice compensation cannot currently be used as an indicator of poor quality of care because of the many variables other than the merits of the case that affect the resolution of individual malpractice court trials and of negotiated settlements. Although it is clear that more and higher payments are made against some specialties than other specialties, there is insufficient evidence to evaluate whether multiple awards against an individual physician indicate poor quality.

Given present information, malpractice litigation information could possibly be used as a screen or trigger for further investigation into a physician's performance by patients, hospitals, liability insurers, and third-party payers. The screen would be weak, since so few people file malpractice claims and resolution often occurs years after the triggering incident (548). Questions of the type of malpractice information (claims, settlements, or jury awards) to be used for screening purposes would need to be decided, as well as how many claims, settlements, and jury awards over what time period would initiate the trigger action.

Before malpractice compensation can be considered an indicator of quality, much more needs to be learned about standards of care. There are disagreements about what constitutes real malpractice, and establishing standards of care might help remedy the problem. Information is needed on the relationship between physician characteristics and medical malpractice claims, judgments and settlements and on physician malpractice profiles and negligently induced adverse outcomes. To understand the relationship between multiple payments and negligence, more needs to be known about the relationship of patient and practice characteristics (e.g., the number of procedures performed) to multiple claims and payments. The Harvard Medical Practice Group is starting to examine medical care and medical injuries in the State of New York. Similar national information is needed on the incidence, severity, and pattern of injuries of negligently induced adverse outcomes. The Harvard group also intends to determine the relationship of adverse outcomes to subsequent tort or disciplinary actions, and the relationship between the probability of suits and the distribution of adverse events and of substandard care.[26]

Government agencies have not traditionally collected data on malpractice. Recently, however, the Health Care Quality Improvement Act of 1986 established a mechanism in Federal law for collection and limited dissemination of information on malpractice payments as well as formal State disciplinary actions, adverse hospital privilege information, and adverse membership actions taken by professional societies. The 1986 act provides that any entity that makes payment under a policy of insurance or self-insurance or in settlement or satisfaction of a judgment in a medical malpractice action or claim must report that information to the Secretary of HHS or the Secretary's designee. The penalty for failure to report malpractice information is a substantial fine. The information that is to be reported includes the physician's name, the amount of payment, and a description of the acts and omissions or injuries upon which the action or claim was based. This information would dramatically improve what is known about malpractice litigation and may offer an opportunity for reexamining the validity of malpractice information as an indicator of the quality of care.

The Health Care Quality Improvement Act may also considerably improve the dissemination of information on malpractice litigation. Currently, dissemination of information on malpractice compensation is limited to information on court awards, which like any other court record is public. The information is published sporadically in costly private newsletters that cover metropolitan areas. The 1986 act requires HHS to make physician-identified information collected in the national data bank available to health care entities and licensing boards. Hospitals are required to obtain the information from HHS, and will be presumed to have the information in any medical malpractice action. Information in the data bank will not be available to individuals. Given the problems of using malpractice compensation as an indicator of the quality of care, publicizing such information to consumers requires further examination.

---

[26] The Robert Wood Johnson Foundation has funded 13 other projects to increase current understanding of what constitutes medical malpractice, what causes it, and how it can be prevented (522).

Combinations of Indicators

A centralized system that includes information on formal disciplinary actions taken by State medical boards, sanctions imposed by HHS upon recommendation of PROs, malpractice compensation, and information on other disciplinary actions taken by medical entities could help to identify recurring problems in the care provided by physicians and perhaps improve the validity of each of the actions as an indicator of quality. Shared information could improve the level of decisionmaking by all concerned bodies. If different, independent bodies censure a physician, the probability that the physician is providing substandard care increases.

A combination of indicators might be a more valid indicator of substandard care than a single indicator. The information could assist in improving future care by making it more difficult than it is now for physicians who have been demonstrated to provide substandard care to continue to practice. However, extreme caution would be needed in using this particular combination of indicators. As discussed above, the validity of medical malpractice claims and compensation as an indicator of the quality of care is not clear. Recent data from the New York State Department of Health indicate that there is a linkage between multiple malpractice claims and disciplinary actions taken by the State medical board (460). Physicians who have had 6 or more medical malpractice claims made against them are likely to be disciplined by the New York State medical board: the State medical board took disciplinary action against 17 percent of such physicians. Further work is needed, since only 181 physicians were studied. The validity of adverse actions taken by hospitals and professional societies also needs to be examined.

The national data bank mandated by the Health Care Quality Improvement Act of 1986 is unique in that malpractice judgments on individuals can be compared with the type of disciplinary actions taken by State medical boards and the adverse actions taken by hospitals and professional societies. Since PRO/HHS sanctions will not be included, the usefulness of the data bank will be limited. Information on such sanctions does not appear to be widely disseminated. The Omnibus Budget Reconciliation Act of 1986 (Public Law 99-509) requires that PROs share, when requested, information related to substandard care with State medical boards and others, but final regulations had not been released by March 1988.

Interest in greater cooperation and sharing of information is seen in the Medicare and Medicaid Patient and Program Protection Act of 1987 (Public Law 110-93). That law strengthens the provisions in the earlier Health Care Quality Improvement Act and requires States to make available to the Secretary of HHS information concerning disciplinary actions taken by State medical boards against a range of health care practitioners. The 1987 law also requires that the Secretary of HHS disseminate information on these actions to State medical licensure boards and to other State and Federal officials.

As noted earlier, information in the data bank mandated by the Health Care Quality Improvement Act will not be available to individuals, and this situation might be reasonable. A prudent course of action in establishing the data bank would be to begin with fairly detailed data but very limited distribution, and then to test the seeming credibility and usefulness of the data as they begin to accumulate for statistical power or actuarial credibility. The data bank will need to be continuously analyzed and revised with continuing experience.

Chapter 7

Evaluation of Physicians' Performance: Care for Hypertension

CONTENTS

Page
Introduction 145
Evaluations of the Outcomes of Care for Hypertension 148
  Reliability of the Indicator 149
  Validity of the Indicator 149
  Feasibility of Using the Indicator 150
Evaluations of the Process of Care for Hypertension 151
  Reliability of the Indicator 152
  Validity of the Indicator 153
  Feasibility of Using the Indicator 156
  Conclusions and Policy Implications 158

Box

Box *Page*
7-A. The Process of Medical Care for Hypertension 146

Figure

Figure *Page*
7-1. The Process of Medical Care for Hypertension 147

Tables

Table *Page*
7-1. Studies on Care for Hypertension Reviewed by OTA 148
7-2. Prognostic Factors for Case-Mix Adjustment of Hypertension Outcome Data 150
7-3. Potential Sources of Information on the Process of Patient Care 152
7-4. Comparison of Generic Approaches to the Evaluation of Physicians' Performance: Care for Hypertension 159

Chapter 7

Evaluation of Physicians' Performance: Care for Hypertension

INTRODUCTION

A major approach to assessing a physician's performance, especially since the 1950s, has entailed evaluation of the care provided for specific medical conditions (184,371). This approach has spread widely during the past two decades as researchers and clinicians have refined assessment techniques. Physicians and other medical professionals have increasingly participated in the review of their peers' performance through privately sponsored activities of hospitals, health maintenance organizations, group practices, medical associations, and third-party payers and through publicly funded programs of State and Federal governments.

This chapter examines the reliability, validity, and feasibility of using evaluations of physicians' performance in caring for a particular condition as an indicator of physician quality. Hypertension is used as a case study condition. Elevated blood pressure is one of the most prevalent and costly medical disorders in the U.S. population, and the effective detection and management of hypertension is one of the Nation's chief public health goals (372,662). Since about 30 percent of the U.S. population is hypertensive,[1] an evaluation of the methods used to assess care for hypertension and to provide information on its quality is important in itself. But evaluating care for hypertension may also illustrate a number of key considerations relevant to evaluating care for other conditions. At the same time, evaluation of the quality of care provided by a physician for hypertension might provide some insights into the quality of other aspects of a physician's services. Consequently, this case study provides a vehicle for analyzing many broader issues in evaluating the process of medical care.

The process of medical care for hypertension is outlined in box 7-A and figure 7-1. In borderline as well as more severe cases, hypertension is generally asymptomatic; its diagnosis depends on the use of blood pressure measurements in individuals who may appear well or who may be seeking care for unrelated health problems. In over 90 percent of cases, hypertension cannot be attributed to an identifiable pathologic cause and must be treated on a chronic, lifetime basis. Detection and followup are crucial, because long-term sequellae of uncontrolled hypertension include serious morbidity associated with strokes, renal disease, cardiac dysfunction, and increased risk of premature death (89). The efficacy of therapies designed to reduce blood pressure toward desired levels in significantly reducing the incidence of these complications was demonstrated in Veterans Administration trials in the early 1970s (676,677). The Hypertension Detection and Follow-Up Program, a 5-year randomized clinical trial with over 10,000 participants, found that a systematic "stepped-care" program for treatment to reduce high blood pressure was associated with significantly higher rates of pressure control and 5-year survival than was "usual" management (313).

OTA's selection of care for hypertension for analysis in this report was based in part on the

---

[1] Estimates of the prevalence of hypertension depend on the precise definition of hypertension adopted. On the basis of the outcome findings of large randomized controlled trials, the Joint National Committee on Hypertension has recommended that patients be diagnosed as hypertensive if the average of blood pressure measurements taken on at least three successive occasions is greater than or equal to 140 mmHg systolic over 90 mmHg diastolic (332). This definition represents a stricter standard than the previous one, which involved repeated measurements above 160/95. Some variation in the specific cutoff pressure levels used for patients in the "mild" hypertensive category still exists among clinicians, especially outside the United States (47). Further, some specialists have argued for diagnosing patients with isolated systolic hypertension as well (717). Because elasticity of the major arteries declines with age, the combined prevalence of isolated systolic and diastolic hypertension for persons aged 65 to 74 is estimated at 64 percent overall and 76 percent in blacks.

Box 7-A.—The Process of Medical Care for Hypertension

Medical care for hypertension—including screening for the disorder and managing therapy for it—is an example of medical care for a specific condition and can be described in terms of the spectrum of medical care presented in chapter 3. There is a high degree of consensus regarding the value of widespread population screening and patient adherence to therapies designed to control elevated blood pressure. Consequently, the basic clinical sequence for effective case finding, diagnosis, and management is especially well defined for hypertension (89,569). This sequence is illustrated in figure 7-1. The figure also notes the many possible stages at which inadequate access to care, discontinuities, and patient dropout can result in care failures and thus poor quality.

Appropriate case-finding procedures are particularly important for two reasons: because general preventive measures for essential hypertension have not been established, and because the disease is both asymptomatic and highly prevalent. The target population for case finding, via standard blood pressure measurements documented at least every several years, is the adult population. Confirming the diagnosis of hypertension requires repeated elevated measurements, taken in different limbs, on each of at least two subsequent visits. This requirement before initiating antihypertensive treatment is a consequence of the frequency of isolated hypertensive readings resulting from stress, daily variations, measurement errors, or other transient causes.

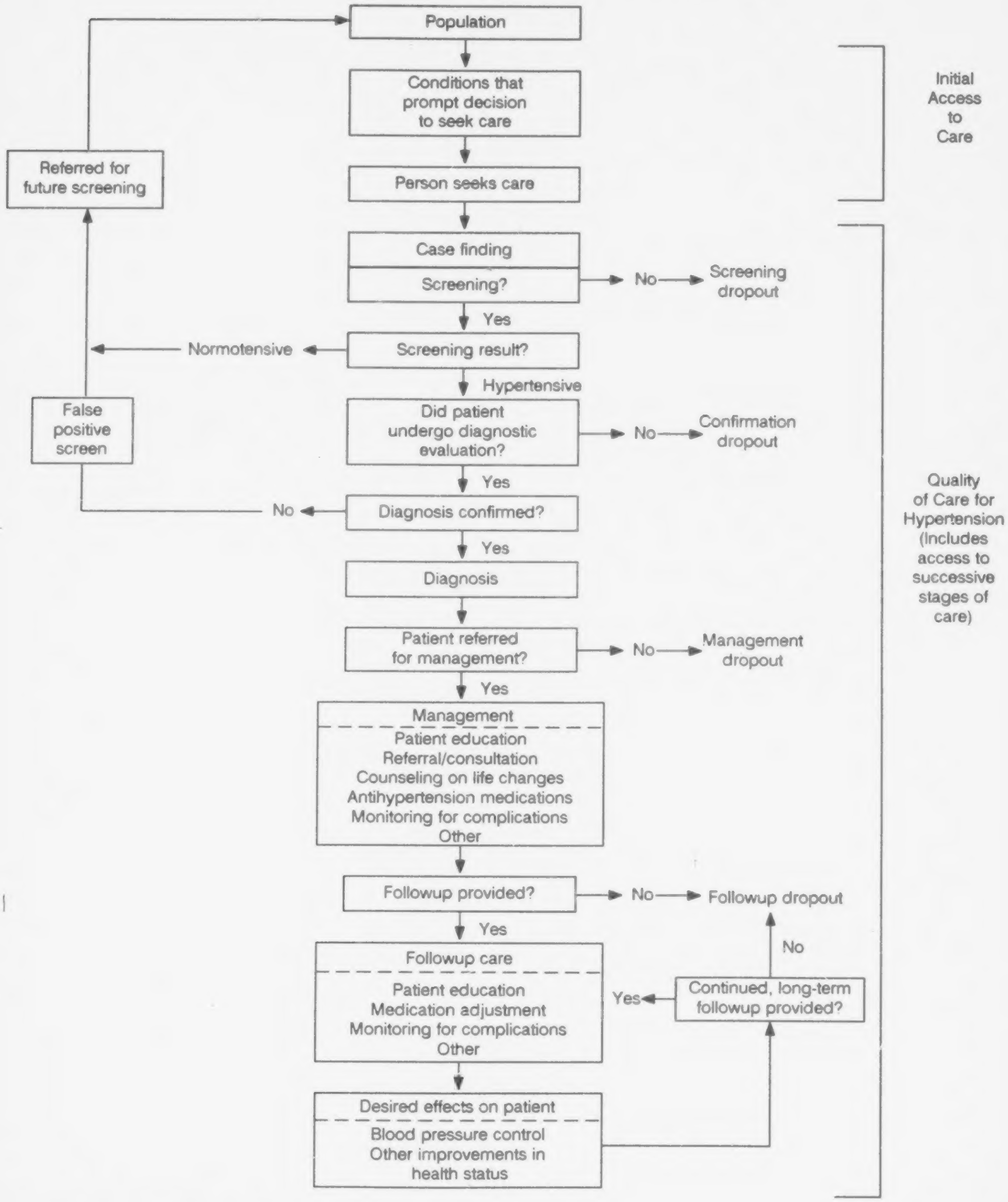
Patients whose diagnosis of hypertension is confirmed represent the target population for management, which involves treatment and followup. Although details may vary among clinicians, treatment typically consists of behavioral modifications coupled with drug therapy. The former include diet modifications to reduce obesity and sodium intake, exercise, cessation of smoking, reduced use of alcohol, and steps to reduce stress, each tailored appropriately to the individual case. Pharmacologic therapy has traditionally featured a "stepped-care" regimen in which more powerful medications are administered incrementally as the patient fails to achieve blood pressure control at a given level (313). These drugs include diuretics, beta blockers, and vasodilators.

The use of stepped-care for certain subgroups of hypertensive patients is currently controversial. These subgroups include patients with mild hypertension (diastolic blood pressure 90-95 mmHg) and patients for whom a particular pathophysiologic mechanism more amenable to an alternative type of medication is suspected (425). The controversy is confined largely to mild hypertensives (and thus is related to controversy in defining hypertension) and to the choice of particular drugs. Broader issues are sufficiently well resolved to permit the elucidation of guidelines for appropriate care.

Because essential hypertension is a chronic condition requiring lifetime treatment, hypertensive patients generally receive care on an ongoing ambulatory basis unless evidence of acute pathological complications supervenes. These complications include strokes, renal disease, visual disorders, or severe headaches. Followup is crucial in management, because patients must adhere consistently to a set of potentially unpleasant behavioral and medical recommendations for many years.

fact that the efficacy of antihypertensive therapy has been well demonstrated and that there exists a fundamental clinical consensus on its effectiveness. The demonstrated efficacy of generally accepted procedures supports the validity of basing quality assessments on adherence to the procedures. Technical aspects of care for hypertension are critical to case finding and management; examples include appropriate screening and diagnostic procedures, proper drug prescriptions, and patient followup for monitoring the effects of therapy and the possible development of complications. It is important to recognize, however, that interpersonal aspects of care for hypertension may be just as important as the technical aspects: hypertensive patients must be persuaded to comply with their medication schedule in spite of unpleasant side effects (196), lifestyle changes may be necessary, and behavior modifications must be maintained. Clearly, both technical and interpersonal aspects of care for hypertension must be considered in evaluating quality. Further, hyper-

Figure 7-1.—The Process of Medical Care for Hypertension


Population
Conditions that prompt decision to seek care
Person seeks care
Initial Access to Care
Referred for future screening
Case finding
Screening?
No
Screening dropout
Yes
Normotensive
Screening result?
Hypertensive
False positive screen
Did patient undergo diagnostic evaluation?
No
Confirmation dropout
Yes
No
Diagnosis confirmed?
Yes
Diagnosis
Quality of Care for Hypertension (Includes access to successive stages of care)
Patient referred for management?
No
Management dropout
Yes
Management
Patient education
Referral/consultation
Counseling on life changes
Antihypertension medications
Monitoring for complications
Other
Followup provided?
No
Followup dropout
Yes
No
Followup care
Patient education
Medication adjustment
Monitoring for complications
Other
Yes
Continued, long-term followup provided?
Desired effects on patient
Blood pressure control
Other improvements in health status

SOURCE: Office of Technology Assessment, 1988

tension is generally diagnosed and managed by a physician in an ambulatory setting rather than in a hospital. Its treatment thus depends on a major segment of health care providers that many of the other potential indicators of quality evaluated in this report do not address.

Drawing on published and unpublished studies (see table 7-1),[2] this chapter analyzes the reliability, validity, and feasibility of using evaluations of physicians' care for hypertension as an indicator of quality. Two generic approaches may be used to evaluate physicians' care:

- evaluations of patient outcomes, and
- evaluations of the process of medical care through the use of explicit criteria or implicit judgment.

The reader should recall that hypertension is only an example and that many of the same concerns that arise may apply to evaluations of physicians' care for other conditions. Clearly, some issues transcend the specific case of evaluating care for hypertension. What are the advantages and disadvantages of using patients' medical records as the source of data for assessments of the process of care? And how can aspects of care that are poorly reflected in medical records best be evaluated? How can physician involvement, and thus medical expertise, be incorporated most effectively into evaluation techniques? How should specific criteria and standards be developed and applied to evaluate physicians' performance? Do evaluations of the process of care need to adjust for differences among patient groups, in disease severity or otherwise? Is the quality of care provided for one condition at all indicative of the overall quality of a physician's practice, or are no such generalizations possible? Most importantly, how can data obtained using these evaluative techniques be appropriately and effectively translated into information useful to consumers? The following analysis discusses these issues in th context of hypertension, but similar issues arise in any attempt to assess physicians' performance by evaluating the care rendered for a specific condition.

Table 7-1.—Studies on Care for Hypertension Reviewed by OTA[a]

Assessments of patients' outcomes:
Brook, 1973 (99)
Schroeder and Donaldson, 1976 (557)
Shorr and Nutting, 1977 (569)
Fletcher, et al., 1979 (211)
Hulka, et al., 1979 (309)
Dove and Schneider, 1980 (188)
Keeler, et al., 1985 (343)

Assessments of medical process, implicit criteria:
Brook, 1973 (99)
Hulka, et al., 1979 (309)
Hastings, et al., 1980 (284)
McAuley and Henderson, 1984 (410)

Assessments of medical process, explicit criteria:
Brook, 1973 (99)
Shorr and Nutting, 1977 (569)
Hulka, et al., 1979 (309)
Deuschle, et al., 1982 (174)
Nutting, et al., 1982 (468)
Sheps and Robertson, 1984 (567)
Borgiel, et al., 1985 (86)
Keeler, et al., 1985 (343)

Combined assessments of patients' outcomes and medical process:
Palmer, 1983 (475)
McCoy, et al., 1988 (417)

[a] Numbers in parentheses refer to numbered entries in the reference list at the end of this report.

SOURCE: Office of Technology Assessment, 1988.

EVALUATIONS OF THE OUTCOMES OF CARE FOR HYPERTENSION

The most widely used measure of patients' outcomes in hypertension studies is a reduction in blood pressure levels or hypertension control rates; this is a proxy measure for longer term clinical complications. Actual measures of complications include specific morbidity rates and mortality differences associated with poor control of hypertension. Few studies of patient outcomes have incorporated functional considerations or other measures related to the patients' quality of life. Further, few studies have based quality-of-care comparisons among different provider groups exclusively on outcome measures.

---

[2] Additional details on the studies reviewed can be found in OTA's technical working paper "Hypertension Screening and Management as an Indicator of Quality: Reliability, Validity, and Feasibility Issues" (415).

Reliability of the Indicator

The procedure used to measure blood pressure is a rapid and accurate procedure when performed by trained personnel. But single measurements of an individual's blood pressure often correlate poorly with that individual's typical blood pressure. Consequently, high false-positive rates (343) and false-negative rates (557) of hypertension have been reported when single measurements are used.[3]

In assessing physicians' performance, quality assessors use blood pressure readings noted in patients' medical records. This approach has the disadvantage of relying on outcome data provided by the physician practice being evaluated rather than by a more independent source (475). Most studies do not provide explicit or quantitative information concerning the reliability of these recorded measurements, because the procedure for measuring blood pressure is technically accurate when performed by qualified personnel and because a series of readings from successive visits is typically reported.

If blood pressure data are grouped into different outcome classes reflecting adequacy of control based on clinical consensus (309), variations in definitions of hypertension may reduce the comparability of results obtained, with identical measurements being categorized differently. Reflecting disagreements among expert judgments, this problem pertains especially to whether diastolic pressures in the borderline 90-95 mmHg range are considered controlled. If such expert classifications are used, a uniform system is required across all providers for reliability.

More innovative approaches to outcome assessment can create special reliability problems. For example, relying on judgments by a panel of experts as to whether a patient's outcome is "improvable" or "unimprovable" requires consideration of the same interrater and interrater reliability issues that arise in process measures (99). Typically, however, such problems arise only in assessment methods using implicit judgments of experts.

Validity of the Indicator

The use of blood pressure readings as a measure of the outcome of care for hypertension is intelligible to average consumers, because such readings are the clinical parameter with which hypertension case finding and management are ultimately concerned.

Even an outcome as immediate as blood pressure readings, however, is the result of a broad range of personal and environmental factors, many of which are beyond the influence of a physician's care. This validity problem can be corrected through standardization of a physician's patient mix based on relevant prognostic factors for desirable or undesirable outcomes. Such methods are analogous to the severity-of-illness adjustments described for patient characteristics in hospital mortality data (see ch. 4). Patient age and other variables that various studies have found to correlate significantly with blood pressure control are listed in table 7-2. Although only one of the studies listed in that table had the statistical power of a large prospective randomized trial (343), the studies collectively indicate that factors as such as the patient's age, race, initial blood pressure, weight, compliance with the prescribed regimen, and access to care can be used to help standardize outcomes across different patient samples.

Although statistical manipulations can increase the validity of the assessment results, a substantial portion of the observed variations in outcomes remains unexplained. Can this portion be attributed exclusively or primarily to the quality of physician care? In general, outcome measures provide little insight into what particular steps a provider may be taking—among a universe of uncontrolled factors in the long-term treatment of a chronic illness—that have a significant impact on the outcomes. This difficulty of attribution is a central problem for any assessment of quality that relies purely on outcomes. Consequently, most assessments use some type of process measure or combine process and outcome approaches (569).

---

[3] For this reason, the clinical diagnosis of hypertension requires elevated pressure recordings on successive visits, and perhaps several readings on each visit.

Table 7-2.—Prognostic Factors for Case-Mix Adjustment of Hypertension Outcome Data

Study[a]	Significant factors
Keeler, et al., 1985 (343)	Initial blood pressure Age Sex Race Location
Dove and Schneider, 1980 (188)	Initial blood pressure Presence of alcohol abuse Weight No treatment in other clinics
Fletcher, et al., 1979 (211)	Lower age Initial blood pressure Patient compliance Prescription of certain medications
Nobrega, et al., 1977 (465)	Initial blood pressure Weight Age

[a] Numbers in parentheses refer to numbered entries in the reference list at the end of this report.

SOURCE: Office of Technology Assessment, 1988.

These conclusions are based on a small number of mostly nonrandom and retrospective studies, a situation that limits analysis of the validity of hypertension outcome measures. The validity of any construct for measuring quality of care depends on the extent to which a statistically significant causal relationship exists between the specific processes performed by the physician and the ultimate patient outcomes observed (185). As in many other areas of quality assessment, additional well-designed studies are required for more powerful conclusions about the use of outcome data. Most importantly, further analyses of external factors that significantly influence observed outcomes for a physician's patients are necessary to develop valid adjustments for severity of illness and other patient characteristics.

Feasibility of Using the Indicator

The advantages of using blood pressure control rates or a related outcome to evaluate the quality of care for hypertension are similar to those of using hospital mortality rates to evaluate hospital care. Both are globally oriented measures, subsuming many aspects of care (and much else as well). With the strong emphasis on outcomes in the general population, these measures are also relatively easy for the public to understand.

But many serious disadvantages accompany these measures. Mortality and morbidity rates for surgical and other inpatient procedures can be computed from data obtained over a relatively bounded time frame (e.g., 30 to 60 days after an operation), but the chronicity of hypertension may require data collection over years for valid assessments of management and control. Extended followup periods present practical methodological problems (557).

Another set of difficulties relates to the feasibility of using patients' medical records.[4] A patient's medical record typically contains the most complete information available on the process of technical care and associated outcomes for patients in both hospitals and ambulatory facilities. It is also the legal record of care, and hospital medical records have been used extensively in evaluations of the quality of inpatient care (185).

The first potential obstacle to medical record review is that medical providers must agree to participate in the review. All experimental assessments have involved voluntary participation, with reported participation rates ranging from 30 percent to over 80 percent. Factors enhancing participation rates include persuasion by colleagues (84) or the involvement of physicians within the practice organization in the assessment process and its treatment as a team effort with constructive goals rather than as an adversarial process (99). Presumably, other incentives or compulsions could also enhance participation. Some studies have noted that physicians who have not been board certified or who are members of smaller practice groups are more likely to refuse to participate; this situation raises questions about the representativeness of the results obtained from these studies (309).

Another group of problems concerns practical issues in collecting data from records for ambulatory patients. Obstacles such as indecipherable handwriting and unretrievable records vary significantly by site and practitioner (479). Neuhaus and colleagues identified three types of difficulties in the data collection process: 1) obtaining a listing of all patient visits by diagnosis, 2) finding charts, and 3) dealing with miscoded or unretrievable records (459). Obtaining a list of visits by diagnosis was impeded by the absence of a uniform method of coding diagnoses, by the fact that practitioners generally did not order their records by diagnosis, and by the need to obtain drug listings from pharmacies in some cases. High miscoding rates may have resulted from clerical recording errors, the listing of a single diagnosis when several were under consideration, and the fact that a hypertension diagnosis may not be confirmed on repeat visits. Neuhaus and colleagues also noted that a pilot study of the office practice being assessed could estimate the amount and type of oversampling required to get an adequate number of "complete" cases for analysis (although these oversampled cases might not be representative).

---

[4] The alternative way to develop a similar data stream is through ongoing independent collection of blood pressure measurements, a method that is expensive and logistically difficult.

Technical progress in the management of data bases and other information systems for recording and retrieving patients' medical records is making such records an increasingly useful source of information on physicians' performance. But as a consequence of current problems, cheap and reliable access to data from all providers remains only a possible goal for the future. Moreover, not only has there been less research using patients' records for ambulatory care than for inpatient care, but also ambulatory records are more likely than inpatient records to be too incomplete to serve as an adequate data source. The consistency of medical record quality tends to be greater for large multiprovider organizations with computerized data bases, but most ambulatory care is delivered in small practices where recordkeeping quality may be much more uneven (475). Blood pressure and some key patient characteristics useful for severity-of-illness adjustments, however, are objective findings that are more likely to be recorded regularly than many details of the medical care process (475).

Although consideration of patient outcomes is obviously an important component of any review of the quality of care for hypertension, relying on blood pressure measurements alone—however easy to abstract from patients' records in comparison to elements related to the process of medical care—would probably require some type of independent auditing mechanism to confirm the accuracy of recorded measurements. Moreover, this approach would not directly encourage better adherence by physicians to effective case finding, diagnosis, and management for all hypertensive patients. That goal requires evaluations of the process of care.

EVALUATIONS OF THE PROCESS OF CARE FOR HYPERTENSION

All evaluations of the process of medical care involve the application of quality standards by experts (184). The types of criteria used in process evaluations span a continuum from purely explicit criteria (completely specified checklists) to purely implicit criteria (unstructured expert analysis). Between these extremes are many possibilities, e.g., the use of explicit guidelines for implicit evaluations by medical experts (284) or the use of a limited set of explicit criteria to target cases likely to be unsatisfactory for implicit review by medical experts (475).[5] To a considerable extent, the strengths and weaknesses of various approaches to evaluating the process of care can be analyzed in terms of trade-offs in reliability, validity, and feasibility along this implicit/explicit continuum.

Any evaluation of the process of care requires a data source that can provide adequate information on the processes used in the delivery of care. Possible sources of information are listed in table 7-3 (475,716). Of the sources listed, only sets of medical records—containing histories of case

---

[5] An alternative approach could measure the percent of patients that complete each state of the treatment sequence (see fig. 7-1). This approach would allow access issues to be incorporated into the assessment (569).

Table 7-3.—Potential Sources of Information on the Process of Patient Care

Sources that rely on data collected by providers	Sources that require independent collection of new data
Medical records	Patient interviews
Prescription records	Patient assessments
Claims forms	Taping/videotaping of patient encounters
Appointment books	Direct observation by experts
Patient tracking systems	"Simulated" patients[a]
Incident reports	

[a] Assessors are trained to give a standardized presentation of a clinical problem and (undetected) to evaluate a physician's management of the condition (716).

SOURCE: Office of Technology Assessment, 1988; adapted from R.H. Palmer, *Ambulatory Health Care Evaluation: Principles and Practice* (Chicago, IL: American Hospital Publishing, Inc., 1983).

management recorded by the health providers involved—are usually detailed enough and accessible enough to be used for evaluating care for specific conditions, such as hypertension.[6]

Reliability of the Indicator

Variations in judgment over time or among physicians represent an obvious problem for a method of quality assessment that uses relatively unstructured expert opinion. Thus, implicit evaluations of the process of medical care must address reliability issues (99). Low interrater reliability may result from systematic bias, with some raters having an inherent tendency to rate cases more stringently than others. These variations can be moderated by adjusting the results statistically to obtain identical mean scores among reviewers (309).

Alternatively, reviewers may simply have different expectations or standards. Various steps can be taken to reduce these interrater differences: selecting physicians who are motivated to participate in the quality review or who are experienced in such assessments, including them in the development of the study, providing clear instructions and guidance, and preparing and distributing case summaries to minimize "nonreviewer" sources of variability (309). Indeed, one observer cites studies indicating that although physicians untrained in abstracting and evaluation have interrater reliability scores approaching 50 percent (the same as pure chance), training physicians in peer review and training abstractors to extract explicit information from records is "reliable and rapid" and results in substantial reliability gains (479). More rigorous studies report that complete agreement among reviewers occurred in 70 to 80 percent of the judgments (99,309,518); less rigorous studies usually obtain higher rates. Findings regarding intrarater reliability have been somewhat more divergent, but generally show slightly higher consistency (e.g., 85 percent in Brook's study).

Richardson concluded that 16 to 28 judges would be required to obtain a reliability of 95 percent for expert evaluation of a given case (518). Brook noted, however, that "unsatisfactory" judgments by two judges indicated that the record involved had a comparable probability of reflecting unsatisfactory care, although only some 20 percent of unsatisfactory cases would be detected (99). Thus, identical judgments by several reviewers may be adequate for detecting unsatisfactory care with a high degree of specificity, but the sensitivity of implicit review methods for identifying particular cases of inadequate care is more questionable.[7]

In explicit evaluations of the process of care, the criteria used in the evaluation are specified in more or less detail, and the reviewer need only determine whether items meeting the criteria are present in the medical record. Consequently, in studies using explicit criteria, high reliability tends to be reported if the reliability issue is addressed at all. A finding well above 90-percent concordance between abstracts by different reviewers, or between staff auditors and project directors, is typical (309,569). More general or nebulous criteria items tend to result in lower reliability (86), and failure to note items present in the record (false negatives) seems more prevalent than crediting items not present (569). Use of physician auditors is not essential for achieving high reliability in explicit evaluations; however, reliability may be significantly enhanced by using reviewers who are familiar with medical terminology and reading medical records (e.g., nurses or graduate students), by providing training sessions, and by conducting periodic reliability checks (309). More limited data on the consistency of a physician's recordkeeping across cases, and thus on the reliability that the provider will consistently record specific process items, yield results that are not quite as impressive but are encouraging (309). Few data on intrarater reliability in explicit evaluations of the process of care are available, probably because interrater reliability is reported to be so high.

---

[6] Of course, medical records have a number of limitations, as described in the preceding section and below.

[7] Specificity and sensitivity are statistical measures relating to accuracy. In this case, specificity represents the proportion of actual cases of satisfactory care that are identified as "satisfactory" (true negative rate). Sensitivity reflects the proportion of actual cases of unsatisfactory care that are identified as "unsatisfactory" (true positive rate). Generally, increasing sensitivity in a measurement system reduces specificity, and vice versa.

The high reliability of evaluations using explicit criteria suggest that steps to make implicit reviews of the process of care more explicit may increase reliability. For example, physician reviewers might be asked to comment explicitly on the basis of their judgments (309). Alternatively, guided criteria for implicit judgments might be developed, such as a checklist to guide reviewers' evaluations of patients' records (284). Even with use of the checklist, however, interrater reliability remained within the range typical for implicit evaluations.

A combination method reported by Palmer used both explicit and implicit approaches (475). This method involved using a small number of straightforward explicit criteria, with which 100-percent compliance was expected. Medical records not in full compliance with these explicit criteria were submitted for implicit judgments on whether the care provided was satisfactory or unsatisfactory. Screening with simple explicit criteria ensures that selection of cases for possible poor-quality care has high reliability.

Explicit evaluations thus have substantial advantages in reliability compared with implicit evaluations, particularly when appropriate steps are taken to promote it. Even though some activities may also enhance the reliability of implicit reviews, these evaluations have substantially less impressive reliability results, especially for the accurate detection of a high proportion of the cases with poor-quality care.

Validity of the Indicator

Despite problems with reliability, review of the process of care by medical experts using implicit criteria is intuitively valid to average consumers, provided that the medical experts involved have acceptable qualifications. The use of explicit criteria have been criticized as invalid because such criteria do not reflect adequately patient heterogeneity and the complexities inherent in clinical practice (309). The use of medical experts theoretically permits clinical insights and consideration of all relevant factors contributing to the management decisions for a specific patient. The severity of a patient's illness, appropriate management of concurrent conditions, and other important elements may be difficult to assess properly with explicit criteria. For this reason, Donabedian concludes that current methods for assessing physician performance using explicit criteria are not substitutes for this comprehensiveness: "For though 'peer review' of the entire record of performance (whether of process alone, or process and outcome combined) is open to error and abuse, as we all recognize, there is nothing we now have that can handle better the entirety of practice in all its rich variety and detail" (184).

Obviously, setting criteria or standards for evaluating the process of care is critical for the authority of an explicit evaluation. Various methods have been used to obtain guidelines applicable to evaluations of care for hypertension. These methods involve variations on either deriving criteria from standards published in the internal medicine literature or developing criteria through some kind of clinical consensus process. One method, for example, involved submitting lengthy questionnaires to two panels of clinicians, generalists, and specialists, and adopting the criteria approved by two-thirds of each group (99).

Other researchers have either developed minimal standards for the various aspects of care (468), relied on criteria developed by an internal physician committee (86,174), used national clinical standards (417), or used items and scoring systems developed in previous process evaluations (567,576). The resulting criteria consequently may reflect guidelines produced or influenced by national or other formal medical organizations, academic physicians, specialists or generalists, or local practitioners. In the most extensive study of the subject, Hulka and colleagues compared results obtained through different criteria-selection mechanisms (309). Even though the lists of criteria and physicians' adherence to them varied, all criteria sets tended to produce parallel results.

Even if relative physician performance using various criteria sets may be similar, criteria lists must be limited not simply to critical items but to critical items likely to be recorded. Patients' medical records emphasize key positive findings, especially objective ones, such as test results. Counseling, communication issues, and other important interpersonal aspects of care are relatively inaccessible to record-based evaluations (475). Further, as Donabedian has noted, critics have argued that the medical record rather than the care itself is being assessed (185). In a review of studies of the validity of the medical record, Hulka and colleagues reported arguments that the legal record of care should be good enough for peer review; they also reported findings that one-third of internists kept records inadequate for review and a study noting poor concordance between written and tape records for information more detailed than a patient's chief complaint and diagnosis (309). In an analysis of the relationship between physician entries and independent records of care, however, Lyons and Payne found that all physician records were complete enough for abstracting and that correlations in adherence scores between the two sets of records were generally significant (400).

Thus, in setting criteria, some tension exists between using a fairly detailed list of evaluative criteria (achieving completeness but emphasizing technical aspects of care and including items more likely to be nonessential, redundant, or unrecorded) and using a shorter, less specific list of criteria (useful for determining if some minimal standards of care have been met) (185). Furthermore, the use of explicit criteria may undesirably reduce physicians' flexibility in approaching the care of a wide range of patients in a wide range of clinical situations or undesirably reduce physicians' incorporation of new clinical knowledge into their practices (185).

These problems and tensions in setting evaluative criteria are well illustrated in a series of studies designed to show a correlation between physician performance in hypertension case-finding and management, as measured by various criteria sets, and patient outcomes. The goal of these studies has been to demonstrate that adherence to criteria lists derived by the methods described above, and presumably reflecting established medical practices, has been associated with favorable patient outcomes. The studies have typically used explicit process measures with the control of diastolic blood pressure as the outcome measure. Several studies have found little or no correlation between process and outcome, even with correction for initial diastolic blood pressure (as an indicator of disease severity) (176,309,339, 465). On the basis of similar results, Romm and Hulka concluded that the setting and promoting of standards for the process of care do not guarantee adequate patient outcomes and that peer review groups should recognize the limitations of both process and outcome measures (530).

Some process-outcome correlations have suffered from poor research designs or statistical analyses, for example, failure to control for patients' initial status in their correlations (411). More fundamentally, many of the evaluative criteria have questionable validity because they are often related to matters such as identifying nonessential causes of hypertension or serious late-stage complications and therefore would not be expected to have a significant impact on overall outcomes. One observer suggests a two-stage approach for the acceptance of specific medical practices as assessment standards: 1) constructing criteria sets based on clinical research concerning diagnostic accuracy and therapeutic effectiveness, and 2) applying the criteria to evaluate physician performance (411).

The key point is that processes believed to have a significant impact on the outcome of care should form the basis for valid assessment criteria (68). Other items, however embedded in customary or established medical practices, should not (476).

For hypertension, examples of processes believed to have a significant impact on outcomes include adherence to a regimen of antihypertensive medications and behavioral and dietary modifications. Patients' knowledge of their disease and adherence to a physician's recommendations for its management appear to depend on the ability of providers to communicate the rationale and benefits of therapy (558). Unfortunately for assessments using medical records, these items all involve key interpersonal components, including patient education and motivation as well as physician discretion.

Methods for measuring these interpersonal aspects of care lack sophistication, but some studies have used rather innovative approaches to address these measurement difficulties and generally have found process-outcome correlations. One study, for example, included a measure of patient compliance with therapy based on the patient's verbal reports about taking prescribed medications, following dietary guidelines, observing recommended changes in activities and habits, and keeping medical appointments (250). Compliance with therapy accounted for a greater portion of the variance in clinical outcomes than the type of therapy, and compliance was also strongly associated with both patient knowledge and perceptions of care. Assessing compliance with the medication regimen by counting the number of pills remaining in patients' prescription bottles, another study found significantly higher rates of blood pressure control among more compliant patients and among patients receiving a more vigorous medication regimen (286). Although patient compliance clearly depends on many factors—psychological, economic, demographic, and other—some of which lie beyond the influence of the physician, these studies indicate that a patient's compliance with therapy has a significant impact on the outcome of care and may be related to the physician's talents in educating, motivating, encouraging continuity of care, and other interpersonal matters.

Just as process and outcome measures may yield divergent results when used to judge the same cases, implicit and explicit process measures may yield results with some divergence (99,309). Implicit ratings for a case tend to be higher than ratings for the same case based on adherence to explicit criteria. Judges using implicit criteria were influenced by favorable outcomes, and they justified their conclusions with items different from the items on the explicit criteria lists; specifically, these judges mentioned procedures related to followup care, criticisms of the physician for performing too many procedures or failing to respond adequately to additional risk factors or comorbid conditions, patient characteristics, and other processes difficult to specify on explicit criteria lists.

Research efforts have led to significant progress in identifying ways to increase the validity of process measures, but a number of difficult issues have not yet been fully resolved. In the validity of process measures, as in the reliability of process measures, trade-offs exist along the spectrum from implicit measures to explicit measures. Because implicit measures allow a patient's medical record to be reviewed in its entirety, they do not break down in the evaluation of cases that are not well suited to a specific set of explicit criteria. Much of the research on process assessment has focused on enhancing the validity of explicit process measures by refining methods for developing and using explicit criteria. Another approach to enhancing the validity of explicit process measures is to combine them with implicit peer review methods. An example is the use of a physician practice audit system that includes a review of each medical record using explicit criteria, which can be performed by nursing personnel, plus a more subjective review performed by a physician (417).[8]

Other validity-related difficulties in assessments of the process of care are common to both implicit and explicit process measures. Both types of measures are limited by the quality of medical records, and neither is well suited at present to evaluating interpersonal and other aspects of care not likely to be found in a patient's medical charts. It is important to note, however, that evaluations of the process of care are the only means of acquiring relatively direct information on whether a physician is following the best clinical practices; outcome assessments cannot be used for this purpose. This fact alone is a very important validity consideration.

In addition to all of the factors related to the validity of implicit reviews and the use of explicit criteria in the evaluation of care for hypertension, another major issue is the extent to which an evaluation of the process of care for hypertension reflects the quality of care a physician is likely

---

[8] Another example is "targeting" implicit review to cases judged unacceptable by simple explicit criteria (475), as noted above.

Photo credit: Harvard Community Health Plan

Combining explicit criteria, such as the monitoring of patients' blood pressure, with experts' implicit judgments improves the validity of using process measures to evaluate physicians' care for hypertension.

to provide for other conditions. Clearly, evaluating care for a single diagnosis appears insufficient to assess a provider's medical abilities generally. Kessner has suggested that the careful selection of a limited set of conditions for evaluation, called "tracers," could provide a framework for evaluating the routine diagnostic, therapeutic, and followup care provided by a health system to the different population groups that it serves (351). Although Kessner was optimistic about the workability of the tracer framework, most subsequent studies purporting to use "tracers" have simply applied the term as a label to the one or several conditions for which care was being evaluated.

There has been little real progress in developing a systematic method to evaluating quality of a physician's care comprehensively with only a limited number of indicator conditions. One study has confirmed the limitations of the generalizability of current explicit performance measures in evaluating internists' management of five hospital diagnoses and six office diagnoses, including hypertension (552). That study found that substandard performance by an internist in managing at least one office condition was associated with a significantly higher proportion of substandard treatment of other office conditions. Substandard office performance by an internist, however, was unrelated to the internist's performance in the hospital, and substandard performance for a hospital condition or superior performance in any condition had no predictive value for substandard or superior performance in other areas. The investigators concluded that the lack of clustering of high or low performance across diagnoses implied that each major diagnostic category in an internist's practice must be assessed independently.

Since a physician's performance in treating one condition does not appear generalizable to the physician's treatment of other diagnoses, an alternative approach is to evaluate a physician's performance across all or most conditions the physician must treat. Borgiel, et al., have developed detailed unweighted explicit criteria sets for 180 conditions most commonly treated by Canadian family physicians (84,85,86). Expanding evaluation to a wide range of diagnoses eliminates the problem of generalizability. But validity issues relating to whether the quality of care for hypertension (or any other condition) can be assessed effectively through outcome or process measures remain.

Feasibility of Using the Indicator

Regarding the feasibility of using evaluations of the process of care to assess quality, the main issue centers on how expert review is incorporated into the evaluation—in developing an evaluative framework (explicit), in the individual reviews (implicit), or in some combination of these stages. Many of the same feasibility obstacles for outcome assessments (plus distinct validity problems) posed by using medical records apply as well.

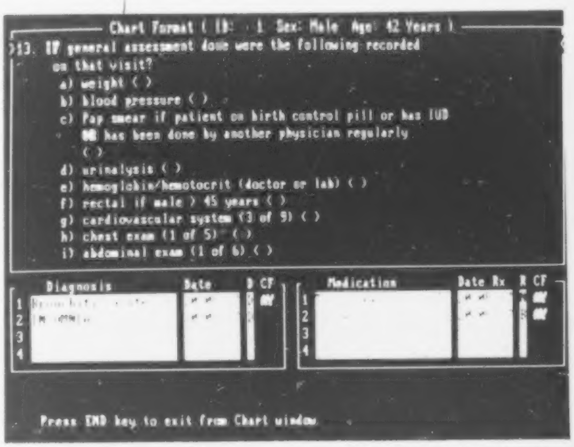
Implicit judgments by medical experts regarding the process of care would have several important advantages in a widespread program of quality assessment to provide information to consumers. Such judgments might be more acceptable to providers as a fair means of assessing the many complex details of individual cases than assessments in which medical professionals do not participate directly (410), a desirable goal since professional support appears to promote the success of an evaluation program (475). Similarly, judgments by clinicians might help promote public confidence in the assessment of a physician's care for a specific condition, since consumers appear to rely heavily on expert opinion in their decisions regarding medical treatment. Further, implicit reviews of the process of care obviate the need for developing and revising criteria lists.

A major disadvantage of implicit assessments of the process of care is their relatively intensive use of expert professional resources. Participation in evaluations would have to become a routine part of the physicians' duties (475), a situation that would reduce their activities in other clinical areas. If such formal responsibilities are not incorporated, record review will probably involve significant delays and inconsistencies (99). These ongoing commitments can be expensive financially as well. The guided implicit review method described by Hulka, et al., required about 15 minutes per case (309) and could be costly (410). Moreover, given the reliability concerns already noted, at least two or three physicians must review each case (and even then high accuracy rates are not guaranteed). Thus, to evaluate a substantial portion of the medical community on a regular basis, an implicit review program would require a major investment of funds and professional time. Additional costs would be incurred for such activities as administration, case abstracting, and training.

In contrast, explicit methods of assessing the process of care have much lower requirements for physician time, since expert participation is limited to developing and revising criteria and reviewing the reliability of the data collection. If training programs are provided, actual record review can be performed by nurses, medical students, and others familiar with the medical environment. The significantly lower expense and higher reliability of explicit reviews may account for their much more frequent use in studies of the quality of ambulatory care. Further, once the criteria and scoring method have been determined, the quantitative data resulting from the analysis can be summarized in a straightforward format to consumers.

These advantages of explicit assessments must be weighed against the validity limitations of such assessments; as noted previously, adherence to criteria lists may not be a fully valid representation of the quality of care provided in specific cases. One likely effect of a policy decision to use explicit criteria to assess the process of medical care would be increased attention to the details of process being measured, possibly at the expense of other aspects of care that might be much more relevant to the clinical outcomes and well-being of a particular patient. Such distortion could be minimized by using only a short list of relatively simple criteria clearly tied to patient outcome, but an assessment based on such a list would probably be capable of determining only whether minimal care was provided.

Borgiel and his colleagues have used explicit criteria to assess the performance of family practitioners (84,85,86). Trained reviewers apply explicit criteria for 180 conditions to review 40 medical records chosen at random. A software program for a portable computer facilitates the abstraction procedure (418). The assessment also includes an interview of the participating physician and a survey of 60 current patients. Each assessment costs about $500 for the patient record audit and $500 for the patient survey (Canadian dollars), costs borne by the physician being assessed. Borgiel and his colleagues recently completed an assessment of 120 family practitioners in southern Ontario. Although participation in



```
------------ Chart Format ( ID:   1  Sex: Male  Age: 42 Years ) ------------
>13. If general assessment done were the following recorded
      on that visit?
       a) weight ( )
       b) blood pressure ( )
       c) Pap smear if patient on birth control pill or has IUD
          OR has been done by another physician regularly
          ( )
       d) urinalysis ( )
       e) hemoglobin/hemotocrit (doctor or lab) ( )
       f) rectal if male > 45 years ( )
       g) cardiovascular system (3 of 9) ( )
       h) chest exam (1 of 5) ( )
       i) abdominal exam (1 of 6) ( )

     Diagnosis           Date     D CF        Medication         Date Rx   R CF
 1                                              1
 2                                              2
 3                                              3
 4                                              4

  Press END key to exit from Chart window
```

Photo credit: College of Family Physicians of Canada

To assess a physician's performance using explicit criteria, trained reviewers use a specialized software program to abstract information from patients' medical records.

the assessment was voluntary, a response rate of over 80 percent was achieved through the use of a recruiting network of clinicians. At present, results are used primarily for educating the assessed physician and for certification decisions by the Canadian College of Family Practice rather than for public information.

Although Borgiel's review of Canadian family practitioners relies exclusively on an explicit method (84,86), other approaches attempt to combine implicit and explicit features with a goal of achieving some of the benefits of each. The targeted method used by Palmer focuses implicit review on cases likely to be unsatisfactory. This method promotes the validity of conclusions about poor quality while reducing expert time—provided, of course, that a high proportion of cases meet the minimal explicit criteria (475). Another example is the explicit/implicit practice audit used in a Minneapolis/St. Paul health maintenance organization (417). Following a phase of feedback and revision to improve the use of the assessment program, this audit system has become regarded as acceptable to most clinicians and is strongly endorsed by the health maintenance organization's management for providing measures of process useful in improving quality of care. The practice audit is expensive, however: the audit requires three nurses and a physician to spend 6 hours on site at the clinic, plus additional time writing the report.

CONCLUSIONS AND POLICY IMPLICATIONS

The most reasonable method for assessing physicians' performance in providing care for a particular condition is to integrate measures of the outcome of care with implicit and explicit measures of the process of care. Depending on the specific method used, a combination of approaches would capture some of the advantages and minimize some of the disadvantages of each generic approach to some extent (see table 7-4). The use of a combined method would be most likely to achieve the goal of promoting reliable and valid judgments as efficiently as possible. Use of combined methods is becoming more common for internal purposes by utilization and quality control peer review organizations (PROs) and by large health care organizations (226), a trend indicating their feasibility.

An effective combined approach could have a range of features, depending on which features of each generic approach to quality assessment are adopted. Cases identified as problematic by the application of specific process or outcome criteria, for example, could be reviewed by physicians, thus providing a check on the validity of the judgment suggested by explicit criteria in a given case (475). Alternatively, physicians using implicit process criteria could review a fraction of the cases randomly selected from a given provider; in the process, reviewers could check whether the results of the explicit evaluation are valid and possibly detect cases of inadequate care that met explicit standards (417). At least at present, some component of peer evaluation appears necessary for supporting the validity of judgments about the adequacy of complex, evolving clinical practices and varied patient characteristics.

Assessment methods that combine the use of explicit criteria and implicit review by medical experts tend to be more expensive than assessment methods based on explicit criteria alone, but combined approaches that target the use of medical experts should cost substantially less than comprehensive peer review systems. The implicit review component of a combined method should be directed primarily toward addressing the weaknesses of the other components of the assessment, such as adjusting for relevant clinical features of the particular case. As assessment methods become more sophisticated, the role of physician review could be refined accordingly, to promote the efficient use of resources in the assessment process.

Although a combined approach to evaluating care for a specific condition appears most promising, many significant obstacles remain for the

Table 7-4.—Comparison of Generic Approaches to the Evaluation of Physicians' Performance: Care for Hypertension[a]

	Reliability		Validity		Feasibility	
Generic approach	+	–	+	–	+	–
Outcome assessment	Blood pressure measurements are accurate	Repeated measurements over time are required; must depend on recording in patient records	Face validity apparent to consumers	*Inadequate sophistication of case-mix adjustment methods Provides no direct information on whether provider is using accepted medical practices	Likely to be recorded in patient records; easy to abstract	*Long followup period likely to be required
Implicit process assessment		Higher intra- and interrater variations because of method's dependence on internal standards (several or more reviewers required)	Face validity for consumers and providers *Theoretically allows comprehensive consideration of all relevant elements in record	Standards may vary or be applied inappropriately	Practice can be audited in a day	*High costs Relatively intensive use of medical professionals
Explicit process assessment	*Specified criteria make measurements easier to replicate		Criteria for judgment are explicit and based on expert standards	Criteria may not fully reflect relevant elements in individual cases Tension between minimal and detailed criteria sets	*Lower cost Less intensive use of physicians *Practice can be audited in a day	

[a] Asterisks are used to designate particularly strong (or weak) features.

SOURCE: Office of Technology Assessment, 1988.

development of any system to assess and disseminate information to consumers about the quality of care provided by individual providers for particular conditions. Some of these problems appear to be organizational and administrative in nature. Additionally, the problems in reliability, validity, and feasibility described in this chapter suggest that important research and implementation tasks remaining before such assessment systems could be realized effectively. Though the chapter has focused on hypertension as a case study, these issues are also relevant to providing information to consumers about the quality of care for other conditions.

Techniques exist to provide such assessments of physician performance. The work of Borgiel and colleagues, McCoy and colleagues, Palmer, and many others indicates that practice assessment systems can be implemented on a continuing basis (84,417,475). Present programs to assess the quality of physicians' care for specific conditions are not designed to provide public information. Instead, they appear to have other worthwhile purposes. Internal quality assessment systems within health care organizations provide feedback and education to providers, to promote quality assurance within a delivery system (58). PROs have been charged with evaluating the quality of ambulatory care of federally funded medical providers; the efforts of PROs are likely to emphasize screening out poor physicians rather than providing consumers information (see ch. 6).

In this institutional context, key administrative and policy issues would have to be resolved before a program could be implemented to provide systematic information about the performance of individual physicians. Since health care providers and organizations do not currently provide such information reliably and in formats useful to consumers, some type of incentive mechanism—either public (e.g., new regulations or enabling statutes) or private (e.g., directives from third-party payers)—would be essential for making the relevant data about patient care available for review. Incentives could be more or less compulsory, ranging from recommendations and voluntary guidelines to requirements that physicians undergo a practice audit as a prerequisite for payment of services, certification, recertification, or licensure. Legal liabilities surrounding peer review and quality assessment would also require analysis (111). As the extant programs indicate, costs for any general audit system would be considerable; they could be borne by the Federal Government or spread among State governments, insurers, other payers, and providers. The issues just cited are only some of the relatively unexamined topics relevant to the successful implementation of a general system of providing physician assessments for consumers.

Moreover, the effects of requirements to disclose information on the dynamics of systems designed for internal quality assurance in health care organizations should be considered. The primary purpose of those systems is to provide effective feedback to improve the quality of work of physicians in the health care organization. But awareness that findings will be made public in a competitive environment could create incentives to minimize the discovery of substandard practices.

In addition to the organizational obstacles cited, many important technical obstacles remain in making this information optimally reliable, valid, and feasible to obtain. Many of these obstacles could be addressed through support from one or more of the research offices in the U.S. Department of Health and Human Services, from the private sector, or from cooperative efforts. To the extent that these difficulties remain unresolved, any assessment method adopted should include features to compensate for the assessment's deficiencies; in this regard, the flexibility provided by combined methods for evaluating the quality of care is especially advantageous.

For evaluations of care using patient outcomes, additional refinements of case mix and severity-of-illness adjustments are needed to make the measures more responsive to the quality of the physician. Additional investigations of methods to increase retention of patients for followup over time and decrease costs of the longer term followup required for adequate outcome assessment of care for a chronic condition might also be useful. For evaluations of care using process measures, investigations of ways to improve the reliability and efficiency of peer review, and in particular applied research on how many physicians and how much of their time is required for a reasonably accurate practice assessment, would permit better use of implicit review methods. Perhaps more importantly, it would be useful to support more sophisticated studies on integrating expert judgments effectively into techniques that also rely in part on adherence to criteria, observed outcomes, or other less costly methods. As a number of investigators have demonstrated, combining features of the different approaches to assessing care can be a very effective way to minimize the weaknesses of individual approaches. A key goal of such studies should be to develop optimal methods in the assessment process for involving physicians, a limited and costly resource.

Another important area for further investigation is determining what relationships exist between the quality of a physician's care for one condition and the quality of the physician's care for other conditions—the issue of generalizability. The few studies that exist provide a sense that each condition is different, but whether assessments can focus on a limited number of diagnoses or must measure the quality of care across the entire spectrum of a physician's practice is obviously a crucial logistical question. Although measures do not appear generalizable at present, more sophisticated analyses might detect underlying patterns or correlations in physician treatment behaviors.

Another key area for further work is the development of better techniques for extracting relevant information from medical records. This is essentially an issue of data quality. Evaluations of care using patients' medical records can assess only items that should be present reliably in the charts, and ambulatory records have much more uneven quality than hospital information systems (479). Increasing computerization of patient data bases is a positive development in this regard. Some larger health care organizations and group practices are relying on such systems, and some quality assessments within hospitals involve manipulation of computerized patient data (446,547). The claims that physicians and hospitals submit to third-party payers could also provide computerized information, especially if entries concerning patients' diagnoses and clinical status were improved. Although a major segment of ambulatory practitioners has not yet adopted computerized office data systems, the creation of some kind of national standards for computerized patient records could be an effective approach to improving reliable access to relevant information on the care process. More generally, uniform standards for data collection and reporting could be developed for all ambulatory records. Such measures would have to consider balancing increased time and cost of more detailed records with the benefits to quality assessment and other activities possible through more reliable or complete data.

Even with such improvements, many critical aspects of medical practice will remain difficult or impossible to capture in a provider's written record. Thus, increasing sophistication in measuring interpersonal aspects of care and physician influence on patient compliance with a therapeutic regimen could result in substantial improvements in the validity of process measures. These deficiencies can be addressed at least in part through patients' assessments of care (see ch. 11), and a physician assessment system featuring medical record reviews complemented by patient surveys could be a powerful approach to developing information on both the technical and interpersonal aspects of care provided. Borgiel and his colleagues currently use this combination in their practice assessments (84,86). Other creative approaches to measuring interpersonal aspects of care, as well as the other physician services not well reflected in the medical record, might also be useful.

Much research has already been devoted to setting standards for evaluating physicians' performance, but the development, evaluation, use, and timely revision of criteria and standards remain a central issue in any assessment that involves explicit criteria. In part, the development of criteria and standards requires clinical studies: much uncertainty remains about what clinical practices and procedures are most strongly associated with medical effectiveness. Ideally, only effective processes should form the basis for criteria developed for evaluations of care (411). In the care of hypertension and some other conditions, the processes that are effective have been relatively well established, and many useful criteria sets have been developed over the last 15 years. Some type of national clearinghouse, perhaps administered through professional medical organizations, might both promote the effective use of these criteria sets and coordinate their refinement with guidelines for the content of medical records.

Improving the quality of ambulatory care assessments will also require further attention to more practical matters related to feasibility. Some of these concerns—such as promoting efficient use of medical experts—have already been mentioned. Many other approaches could also lead to lower assessment costs; examples include improved training methods, improved coordination with other quality-related projects and with organizations and activities designed to promote medical quality, and innovative approaches such as self-audits (417). Another key area is the adaptation of computer technologies to assist in the collection of assessment information. For example, office audits can be expedited using software programs to enter data on adherence to criteria (419). Conceivably, these methods could be coordinated with computerized data base record systems to make assessments more fully automated.

Two other crucial considerations related to the feasibility of using evaluations of physicians' management of specific conditions to evaluate quality deserve final mention. One is the need for further deliberation on whether attention to all of the research items detailed above is worthwhile, or whether less ideal or entirely different approaches would be better alternatives for providing consumer information or for increasing the likelihood that patients will receive high-quality care for specific conditions, such as hypertension. Although considerable experience with assessment methods in both research and practical settings indicates that these methods—especially combined approaches—have considerable promise, the discussion in this chapter suggests that serious technical, organizational, and economic obstacles remain before a functional system could be implemented nationally to provide useful information to consumers about individual physician performance for certain conditions. In this regard, it is important to recognize that almost no research has been directed specifically toward the question of providing information to consumers about the quality of the processes of care they receive for the treatment of hypertension or any other condition.

The other crucial consideration, running through out this chapter, is that evaluations of the process of care clearly require the leadership and assistance of the medical profession. Historically, professional medical associations have played the paramount role in evaluating physicians' performance; at present, they are continuing to expand their activities in promoting high-quality care. Independently of its own assessment activities, or in coordination with them, the Federal Government can support the medical profession's efforts.

Chapter 8

Volume of Services in Hospitals or Performed by Physicians

CONTENTS

Page
Introduction 165
Reliability of the Indicator 166
Validity of the Indicator 167
  Measures of Volume and Outcome 167
  Differences in Patient Characteristics 169
  Research Findings 170
Feasibility of Using the Indicator 177
Conclusions and Policy Implications 180

Figures

Figure *Page*
8-1. Hypothesized Relationship Between Volume and Outcome 165
8-2. Ratio of Actual to Expected Mortality Rates by Volume of Patients Undergoing Coronary Artery Bypass Graft Surgery in California, 1983 170
8-3. Number of Studies Reviewed by OTA Showing Either Worse Outcomes at Low Hospital Volume or No Effect, by Diagnosis or Procedure 175
8-4. Comparison of Actual and Expected Mortality Rates for Patients Undergoing Coronary Artery Bypass Graft Surgery in California, 1983 179

Tables

Table *Page*
8-1. Studies Reviewed by OTA on the Relationship Between Volume and Outcome for Specific Diagnoses and Procedures 168
8-2. The Hospital-Volume/Outcome Relationship: Summary of Research Findings From Studies Reviewed by OTA on Specific Diagnoses or Procedures 172
8-3. The Physician-Volume/Outcome Relationship: Summary of Research Findings From Studies Reviewed by OTA on Specific Diagnoses or Procedures 173

Chapter 8

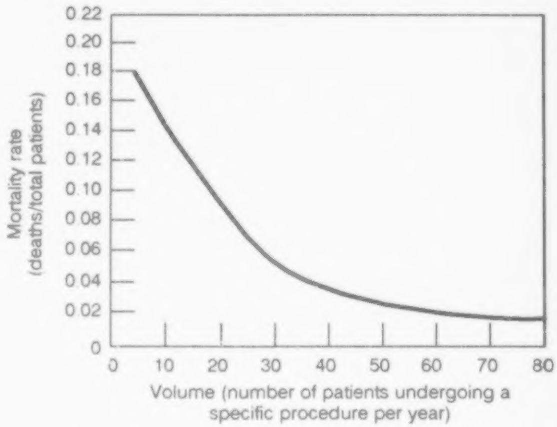
Volume of Services in Hospitals or Performed by Physicians[1]

INTRODUCTION

There is a common notion that "practice makes perfect." In the medical care setting, this adage is often interpreted as "high-volume hospitals and physicians achieve better outcomes." The word "volume" in this context refers to the number of procedures or number of patients with the same diagnosis treated in a specific hospital or by a particular physician. For some procedures and diagnoses, better patient outcomes and lower inhospital mortality have been associated with higher volumes.

In its simplest form, the hypothesized relationship between volume and outcome may be displayed as a graph with volume (e.g., number of patients undergoing a specific procedure per year in a hospital) on the horizontal axis and outcomes (e.g., mortality rate) on the vertical axis. The graph in figure 8-1 shows high mortality in hospitals with low volumes and low mortality in hospitals with high volumes. The flattening of the curve at high-volume levels indicates that there is little additional reduction in mortality above a certain volume threshold.

Figure 8-1.—Hypothesized Relationship Between Volume and Outcome



SOURCE: Office of Technology Assessment, 1988

It is important to limit the conclusions drawn from this graph. Even if a relationship is found between volume and outcome, it is inappropriate to conclude that increasing the volume in a hospital will improve outcomes or that reducing the volume will worsen outcomes. Conclusions cannot be drawn about how changes in volumes affect changes in outcomes, because most analyses use data from a cross section of hospitals observed at a point in time rather than data from the history of mortality and volume over time. Instead of causality from volume to outcome, there may be causality from outcome to volume; that is, medical providers with low mortality rates may attract higher volumes of patients. Another possibility is that some unmeasured factor may account for an observed relationship between volume and outcome. For example, high-volume hospitals or physicians may have relaxed admission criteria; these relaxed criteria, in turn, may mean that some of their patients are healthier and less likely to suffer adverse outcomes. In this case, both the higher volume and the better outcomes may be caused by the relaxed admission criteria.

For OTA's review of the literature on the volume-outcome relationship, the abstracts of approximately 100 papers were read. Of the 50 articles that were thoroughly reviewed, 26 presented reportable findings.[2] Studies were included if they examined a sufficient number of hospitals (over 20) and cases to offer statistically valid volume-outcome results *or* if the study purported to examine the volume-outcome relationship. These studies pertain to both hospital and physician volume. Although most research relates to hospital volume, a growing body of literature focuses on the relative importance of physician volume in contrast or in addition to hospital volume.

---

[1] This chapter is based on a paper prepared for OTA by Harold S. Luft, Deborah W. Garnick, David Mark, Stephen J. McPhee, and Janice Tetreault (395).

[2] Additional technical information on the studies included in the literature review is available in the paper on volume prepared by Luft and colleagues (395).

This chapter examines the reliability, validity, and feasibility of using the volume-outcome relationship as an indicator of the quality of medical care, and explores the issues of causality as well as other relevant conceptual and methodological issues. How volume data might be used by consumers in choosing hospitals and physicians is discussed, and further necessary research is outlined.

RELIABILITY OF THE INDICATOR

Information on the volume of procedures and diagnoses and on inhospital mortality is routinely available from two sources: hospital discharge abstracts and insurance claims. There are several problems with the reliability of hospital discharge abstract data. Errors can occur at different points during the data collection process: in recording the patient's diagnosis or procedure onto the medical chart, in the translation of the chart onto discharge abstract forms, or in the transformation of discharge abstract forms into large-scale computerized data systems. Several studies of the accuracy of hospital abstracting suggest a high error rate (450,532). Moreover, inaccuracy in the data may be the result of random errors, such as misapplication of coding rules or the selection of vague diagnosis codes, or may be the result of purposeful misspecification of a patient's principal diagnosis in order to achieve an optimal diagnosis-related group (DRG) for Medicare payment purposes. Recently, a reabstracting study noted that incorrect DRGs were originally assigned 20.8 percent of the time in 1984-85 and that 61.7 percent of these errors benefited the hospital (304).

Insurance claims data—especially non-Medicare insurance claims data—usually include less information about diagnoses than do routinely collected hospital discharge abstract data. Moreover, coding problems in the case of claims data may be worse than those in the case of hospital discharge abstracts. The problem is especially acute for diagnoses; procedures are generally well coded (131).

The pertinent question here, however, is not whether coding errors occur, but how such errors affect volume-outcome studies. The miscoding of a diagnosis or procedure may cause undercounts or overcounts of the number of patients in certain categories. Many of the diagnoses and procedures that have been studied in the volume-outcome literature are so important to a patient's hospitalization and the categories are so broad, however, that miscounts of patients are probably not an important concern. Total hip replacement, for example, would be unlikely to be overlooked. Moreover, in many studies, volume is specified as a series of categories (e.g., high, medium, and low), so a small amount of random undercounting or overcounting is not crucial. Miscoding of patients' simultaneously existing illnesses (comorbidities) may be a problem in case-mix adjustments to reflect patient differences. The problems in adjusting for patient differences in the analysis of volume are similar to those present in the analysis of hospital-specific mortality outcomes (see ch. 4).

Volume-outcome studies are generally cross-sectional, and changes in the accuracy of data over time are less important than systematic differences across hospitals. When analyses are focused on individual hospitals, the reliability of data is an important concern, because misclassification could result in the mislabeling of a hospital as a good- or poor-quality provider. When the investigation concerns the identification and exploration of the hypothesized relationship between volume and outcome, the reliability of data is less of a key concern. Suppose there are random errors across hospitals in the coding of diagnoses. Such errors will affect the precision with which relationships are estimated, but if the errors are uncorrelated with volume, the volume-outcome effect will not be altered.

VALIDITY OF THE INDICATOR

Table 8-1 presents a summary listing of the 15 procedures and diagnoses investigated in the 26 studies used for the analysis in this chapter. The studies are grouped in the left hand column by research team and by the publication date of the first article by the team (e.g., all three studies by Kelly and her colleagues are shown together). To check which authors studied a particular diagnosis or procedure, read down the column for a given procedure or diagnosis.

Of the 15 procedures and diagnoses investigated in the 26 studies, 13 are surgical procedures. Only 2 are medical diagnoses: acute myocardial infarction ("heart attack") and newborn diseases. The study of surgical procedures is easier than the study of medical diagnoses for several reasons. First, surgical procedures are generally well identified and coded both on hospital discharge abstracts and insurance claims. The occurrence of an operation is rarely in dispute, even though the choice of procedure or necessity for it may be questioned by various physicians.[3] The determination of some diagnoses, on the other hand, is often quite difficult; comparably trained clinicians may disagree on an individual patient's diagnosis.

Second, although severity of illness may vary with both surgically treated patients and medically treated patients, it is less likely to be a major source of bias in volume-outcome studies of patients treated surgically. Surgery is usually used to increase longevity or to correct a problem that interferes with the quality of a person's life but is not immediately life-threatening. Thus, a surgically treated patient is often in reasonably good health on admission to the hospital, and short-term mortality is more likely to reflect the effects of treatment than to reflect the patient's initial health status. In medical admissions, on the other hand, there is greater variation in the complexity of cases, and a patient's health status on admission may be a more important determinant of short-term outcomes than the quality of care rendered is. Thus, the paucity of good measures of patients' severity of illness probably has a greater impact on studies involving medical admissions than studies involving surgical admissions.

Measures of Volume and Outcome

Volume is measured in several ways in the 26 studies reviewed by OTA:

- categorical variables (e.g., low-and high-volume groups, or a four-or five-category classification),
- a continuous variable (e.g., a count of number of patients, which allows for a linear relation),
- volume and volume squared (which allows for either linear or "U"-shaped curves), or
- log of volume (which allows for a stronger effect at low volumes and progressively weaker effects at higher volumes).

Most of the studies measure volume for a single year, although some studies use other periods. One study uses a hybrid: the proportion of patients in a hospital (a continuous measure) treated by surgeons with low volumes (a dichotomous variable) (307).

Four measures of patient outcomes are used in the 26 studies:

- inhospital mortality,
- mortality within a fixed period of time,
- complications or health status measures, and
- long hospital stays as a proxy for complications.

The use of mortality as an outcome measure of quality has some limitations (see ch. 4). For some procedures and diagnoses, mortality is so rare an event that it is difficult to determine whether an occasional death indicates a pattern of poor quality or a chance occurrence. Important biases may also be introduced because discharge policies controlled by hospitals can affect inpatient mortality rates. In hospitals that transfer patients with severe complications to other, more appropriate facilities, such as regional tertiary hospitals, there are likely to be lower mor-

---

[3] In some cases, there may be miscoding of which procedure occurred, for example, revision of total hip versus total hip replacement.

Table 8-1.—Studies Reviewed by OTA on the Relationship Between Volume and Outcome for Specific Diagnoses and Procedures

	Abdominal aortic aneurysm	Acute myocardial infarction	Appendectomy	Biliary tract surgery	Cardiac catheterization	Coronary artery bypass graft surgery	Femur fracture	Hernia	Hysterectomy	Intestinal operation	Newborn diseases	Prostatectomy	Stomach operation	Total hip replacement	Vascular surgery
1. Adams, et al., 1973 (5)					HV/D										
2. Williams, 1979 (702)											HV/D				
3. Luft, et al., 1979[b] (394)	HV/D			HV/D		HV/D				HV/D		HV/D	HV/D	HV/D	HV/D
4. Luft, 1980[c] (393)	HV/D			HV/D		HV/D				HV/D		HV/D	HV/D	HV/D	HV/D
5. Maerki, et al., 1986[d] (402)	HV/D	HV/D	HV/D	HV/D	HV/D	HV/D	HV/D	HV/D	HV/D	HV/D	HV/D	HV/D	HV/D	HV/D	
6. Luft and Hunt, 1986 (396)					HV/D										
7. Luft, et al., 1987[e] (397)	HV/D	HV/D	HV/D	HV/D	HV/D	HV/D	HV/D	HV/D	HV/D	HV/D	HV/D	HV/D	HV/D	HV/D	
8. Hughes, et al., 1987 (307)			HV,PV/D,L	HV,PV/D,L	HV,PV/D,L	HV,PV/D,L		HV,PV/D,L	HV,PV/D,L	HV,PV/D,L		HV,PV/D,L	HV,PV/D,L	HV,PV/D,L	
9. Hughes, et al., in press (306)							HV/D,L								
10. Pilcher, et al., 1980 (488)	HV,PV/D														
11. Farber, et al., 1981[f] (203)			HV/M					HV/M	HV/M	HV/M					
12. Shortell and LoGerfo, 1981 (571)		PV/D	PV/M												
13. Hertzer, et al., 1984 (295)															PV/D,M
14. Flood, et al., 1984[g] (217)	HV/D			HV/D			HV/D			HV/D			HV/D		
15. Rosenblatt, et al., 1985 (538)											HV/D				
16. Riley and Lubitz, 1985 (520)				HV/D		HV/D	HV/D	HV/D		HV/D		HV/D		HV/D	
17. Kempczinski, et al., 1986 (349)															PV/D,M
18. Sloan, et al., 1986[h] (582)						HV/D			HV/D					HV/D	
19. Kelly and Hellinger, 1986 (347)	HV,PV/D									HV,PV/D			HV,PV/D		
20. Kelly and Hellinger, 1987 (348)		HV,PV/D			HV,PV/D	HV,PV/D									
21. Kelly, forthcoming[i] (346)				HV/D	HV/D	HV/D									
22. Roos, et al., 1986 (531)				HV,PV/R					HV,PV/R			HV,PV/R			
23. Roos, et al., 1987 (533)				HV/R					HV/R						
24. Wennberg, et al., 1987 (697)												HV/D,R			
25. Showstack, et al., 1987 (573)						HV/D,L									
26. Fowles, et al., 1987 (227)														HV,PV/M,D	

Abbreviations: HV = hospital volume; PV = physician volume; D = death; L = long length of hospital stay; M = morbidity; R = readmission.

[a] Studies are ordered by research team and date of first publication by the team. Numbers in parentheses refer to numbered entries in the reference list at the end of this report.

[b] Luft, et al. (1979), also studied open-heart surgery.

[c] Luft, et al. (1980), also studied open-heart surgery.

[d] Maerki, et al. (1986), also studied cirrhosis, peptic ulcer, subarachnoid hemorrhage, and tonsillectomy.

[e] Luft, et al. (1987), also studied cirrhosis, subarachnoid hemorrhage, and peptic ulcer.

[f] Farber, et al. (1981), also studied laminectomy and cesarean section.

[g] Flood, et al. (1984), also studied amputation of lower limb, nonsurgical gallbladder diagnosis, and nonsurgical ulcer diagnosis.

[h] Sloan, et al. (1986), also studied morbid obesity surgery, mastectomy, nephrectomy, and spinal fusion.

[i] Kelly (forthcoming) also studied atherosclerosis, cranial injury, diabetes, and hypertension.

SOURCE: Office of Technology Assessment, 1988.

tality rates. In hospitals with longer average stays, there is a greater chance of observing a death. Suppose, for example, that one hospital typically keeps patients for 10 days after a certain surgical procedure, while another hospital works to get patients on their feet and discharges them after a week. If a certain fraction of patients from each hospital experiences a fatal heart attack on the 8th to 10th days after surgery, these deaths will be counted in the inhospital mortality rate of only the first hospital. Because of these biases, some researchers calculate mortality rates with respect to a fixed window, such as 30 days, after admission (643).

Complications and other measures of patients' health status are less objectively measured than mortality. In some instances, a clearly identified procedure, such as a reoperation, indicates a poor outcome. Other measures, such as surgical wound infections, are less reliably coded across hospitals (see ch. 5).

One final measure of quality is even further removed from a direct measure of outcome. Luft and his colleagues use the proportion of patients that stay a very long time in the hospital as a proxy for complication rates (306,307,573). They argue that if one chooses a length of stay exceeded by only 10 percent of all patients, then a hospital with far more than 10 percent of its patients staying that long or longer may be experiencing poor outcomes. Although this argument is plausible, it has not been validated by determining whether those patients with very long hospital stays truly have complications, or stay longer, for example, because nursing home beds are scarce.

Differences in Patient Characteristics

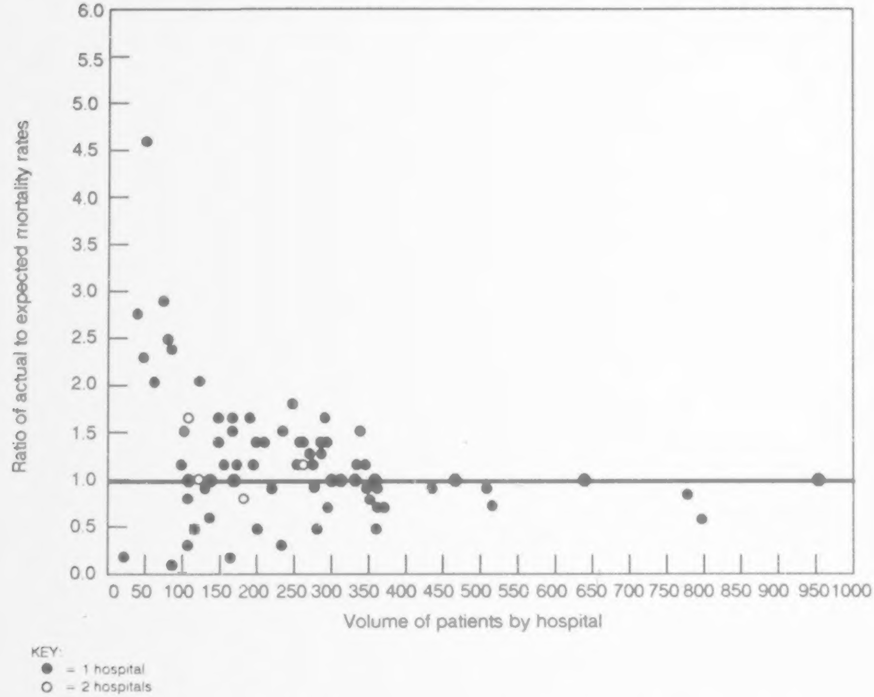
A major problem in analysis of the volume-outcome relationship is the potentially confounding effect of differences in patient characteristics. Every patient is different, and individual factors strongly influence outcomes. Even if these patient differences are random, the estimation of a volume-outcome relation will be made more difficult because of the "noise" due to these random effects. This point is illustrated in figure 8-2 which plots the inpatient mortality rates for patients undergoing coronary artery bypass graft (CABG) surgery in 78 California hospitals in 1983 (574). Although there generally appear to be lower rates of poor outcomes at higher volumes (a negative linear relationship), there is substantial variation among hospitals at given volume levels—variation due in part to patient-related factors.

The crucial question is whether more or less severely ill patients are consistently admitted to high-volume hospitals. If they are, an observed association between outcome and volume could be due entirely to patient mix. The true answer to this question would be found by random assignment of large numbers of patients to institutions with varying volume levels. Random assignment, with sufficiently large numbers of patients, would reduce to insignificance the likelihood that patient-related factors account for the observed differences in outcomes. Unfortunately, since such an experiment would be enormously expensive and impossible because of ethical considerations, one is left with attempts to control for patients' differences by various statistical means.

There are two general approaches to dealing with differences in patient mix across hospitals. The first is to specify the procedure or diagnosis for study as carefully and narrowly as possible. The intent of this approach is to set patient selection criteria that result in a homogeneous group of patients. For example, patients undergoing CABG surgery who also have heart valve surgery have mortality rates about three times as high as those of patients undergoing CABG surgery only (573). Since some hospitals may specialize in uncomplicated CABG surgery while others have a large share of patients also requiring valve surgery, results may be biased unless one focuses on patients with CABG surgery only.

The second approach, which can be combined with the first, is to include variables in the analysis that may capture risk differences among the patients included in the study. In theory, each of these additional variables could be used to further stratify the study population of patients, but this approach is limited by an ever shrinking sample size. In many studies, therefore, patient selection criteria are combined with statistical controls. The patient's age, race, and sex are classic variables used in analyses. Transfer from another hos-

Figure 8-2.—Ratio of Actual to Expected Mortality Rates by Volume of Patients Undergoing Coronary Artery Bypass Graft Surgery in California, 1983



SOURCE: J.A. Showstack, K.E. Rosenfeld, D.W. Garnick, et al., Institute for Health Policy Studies, University of California, unpublished data, San Francisco, 1987

pital is often a powerful indicator of a patient at higher risk of a poor outcome (393). Counts of the number of secondary diagnoses or procedures or the presence of specific diagnoses or procedures also are used (394,573,582). In some instances, diagnostic information is combined to form a disease "stage" indicative of the severity of the principal diagnosis (346,347,348).

The problem of differences among patients has been highlighted in the literature on using mortality data to evaluate hospital performance (78, 189). To some extent, the problem is more severe if the focus is on studying individual hospitals rather than on studying the hypothesized relationship between volume and mortality. If a specific hospital is identified as having a significantly above average mortality rate, the hospital administration is likely to claim that unmeasured differences in patient mix account for the observed results. Upon careful examination of the medical records, one may find that some patients entering the hospital with severe problems do account for an elevated mortality rate (see ch. 4). Precisely what clinical characteristics, if any, are similarly correlated with volume is not clear.

Research Findings

Statistical methods used in the volume-outcome studies listed in table 8-1 range from simple comparisons of high- and low-volume groups to fairly sophisticated causal models. Regression models were commonly used because they can include a large number of patient and/or hospital variables as explanatory factors. In some cases, logistic models were used to account explicitly for the 0,1 nature of patient mortality. Three papers used simultaneous equation models to estimate both the influence of volume on outcomes and the influence of outcomes on volume (307,393,397).

Some researchers used the patient as the unit of observation in a regression model to "predict" the patient's outcome. These researchers included as many patient risk factors as possible as well as variables indicating the number of patients with the procedure or diagnosis in the hospital per year. Patient-level regressions typically produced very low R-squares, indicating a low ability to predict whether an individual patient will live or die, even though many of the variables, such as volume, may be highly significant. Other researchers have argued that the focus of volume-outcome studies is on the average performance of hospitals at different volume levels, so the number of observations should be the number of hospitals rather than the number of patients undergoing procedures (701). These researchers estimated models at the hospital level that include the proportion of patients with each risk factor to predict a composite expected poor outcome rate based upon patient mix in the hospital.

It is difficult to combine or compare results across studies in a formal manner because of the differences in methods. For example, it is impossible to compare directly the findings of one study that simply contrasts mortality rates for hospitals with volumes above or below an arbitrary cutoff with another study that estimates the influence of the log of volume on outcomes while controlling for numerous hospital characteristics and referral effects.

To overcome this problem, OTA categorized the results of each study in terms of the implicit shape of the volume-outcome relationship "curve." Tables 8-2 and 8-3 summarize the results of this categorization for hospital and physician volume, respectively. Potential categories illustrated by the curves in the far left column of each table are as follows:

1. dichotomous results, with volume grouped into two categories and results showing lower rates of poor outcomes in high-volume settings;
2. a negative linear relationship, also showing lower rates of poor outcomes in high-volume settings;
3. a "U-shaped" relationship showing higher rates of poor outcomes at lower volumes, lower rates of poor outcomes at intermediate volumes, and higher rates of poor outcomes at higher volumes;
4. an inverse logarithmic form, with large reductions in the rates of poor outcomes as low volumes increase and a relative flattening at high volumes;
5. a "flat" curve, indicating no significant relationship; and
6. a positive linear relationship, with higher rates of poor outcomes at higher volumes.[4]

Tables 8-2 and 8-3 should be read along with table 8-1, which lists the 26 studies included in OTA's literature review along with the diagnoses and procedures examined in each study. In table 8-2, for example, results for abdominal aortic aneurysm are shown in the first column. The first study in that column, number 14, refers to study number 14—Flood, et al., 1984—listed in table 8-1.

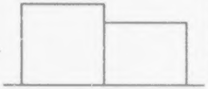
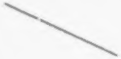
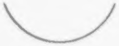
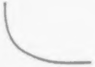
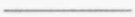
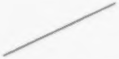
When one of the 26 studies includes two methods (e.g., dichotomous and continuous-volume variables) or differentiates between two subcategories of a procedure (e.g., ruptured aneurysm surgery and elective aneurysm surgery), the results are counted separately. As an example, table 8-2 shows that seven studies addressed hospital volume and abdominal aortic aneurysm. Using regression analysis, Luft and his colleagues found an insignificant (flat) relationship between volume and outcome for this procedure (study 7c); however, using volume categories (study 7a), the same authors found a negative linear relationship between volume and outcome for this procedure (397). Table 8-2 shows both these results.

When reviewing a set of findings such as these, which are relatively thin but nonetheless cover many procedures and diagnoses, one is torn between a "lumping" approach to provide a gestalt and a "splitting" approach to explain differences. Several points can be highlighted for specific diagnoses or procedures. For biliary tract surgery, it is important to distinguish the type of surgery, because the volume-outcome relationship may be valid only for more complex surgery that com-

---

[4] One set of findings by Sloan, Perrin, and Valvona (582) approximates a backwards "C" and is not classified in this schema. These investigators' other findings fit a "U"-shaped pattern and are included in the table.

Table 8-2.—The Hospital-Volume/Outcome Relationship: Summary of Research Findings From Studies Reviewed by OTA on Specific Diagnoses or Procedures*

Shape of the curve	Abdominal aortic aneurysm	Acute myocardial infarction	Appendectomy	Biliary tract surgery	Cardiac catheterization	Coronary artery bypass graft surgery	Femur fracture	Hernia	Hysterectomy	Intestinal operation	Newborn diseases	Prostatectomy	Stomach operation	Total hip replacement	Vascular surgery
1. 	14			14,23					23	14					
2. 	5,7a, 10b	7a	5,7a	5,7a	1,5	7a,25h			5,7a	5,7a		7a	7a	7a	3
3. 		5			7a	5,18		5			2,5	5	5	5,18	
4. 	3,4,19	7c	8,11	3e,3f,4e, 4f,4g,7c, 11	6,8,20, 21	3,4,8, 16,20, 25i,25j	9l,9m	7a,8,11	7c,8,11	3,4,7c, 8,11,16	7a,7c	3,4,8,16	3n	3,4,7c, 8,16r	4
5. 	7c,10d	20	7c	3g,8,16, 21,22		7c,21, 25k	5,7a,7c, 14,16	7c,16	18,22	19		7c,22,24	3o,4n, 4o,7c,8, 14,19p, 19q	16s,26t, 26u	
6. 											15				

*Note: The curves in this table illustrate the following relationships between volume and outcome:

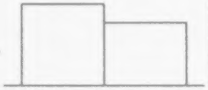
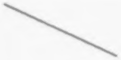
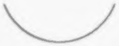
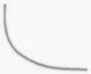
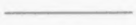
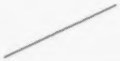
1 Volume grouped into two categories, with lower rates of poor outcomes at higher volumes,
2 Downward-sloping line (negative linear relationship), with lower rates of poor outcomes at higher volumes,
3 U-shaped curve, with higher rates of poor outcomes at lower volumes, lower rates of poor outcomes in intermediate volume ranges, and higher rates of poor outcomes at higher volumes,
4 L-shaped curve (inverse logarithmic form), with large reductions in rates of poor outcomes as volume increases from low levels and little change at higher volumes,
5 Flat, indicating no significant relationship between volume and outcome, and
6 Upward-sloping line (positive linear relationship), with higher rates of poor outcomes at higher volumes

The numbers in the entries for this table refer to the numbered references in table 8-1 The meanings of the letters in the entries are as follows:

a Analysis using volume categories
b Ruptured aneurysm surgery
c Analysis using regressions
d Elective aneurysm surgery
e Cholecystectomy with common bile duct exploration
f Other biliary tract surgery
g Cholecystectomy alone
h Nonscheduled CABG, death outcome
i Scheduled CABG, poor outcome
j Nonscheduled CABG, poor outcome
k Scheduled CABG, death outcome
l Femur fracture, death outcome
m Femur fracture, poor outcome
n Vagotomy and/or pyloroplasty for duodenal ulcer
o Vagotomy, all
p Stomach operations, cancer diagnosis
q Stomach operations, ulcer diagnosis
r Other hip arthoplasty
s Total hip replacement
t Total hip replacement, death outcome
u Total hip replacement, major complications

SOURCE: Office of Technology Assessment, 1988

Table 8-3.—The Physician-Volume/Outcome Relationship: Summary of Research Findings From Studies Reviewed by OTA on Specific Diagnoses or Procedures*

Shape of the curve	Abdominal aortic aneurysm	Acute myocardial infarction	Appendectomy	Biliary tract surgery	Cardiac catheterization	Coronary artery bypass surgery	Hernia	Hysterectomy	Intestinal operation	Prostatectomy	Stomach operation	Total hip replacement	Vascular surgery
1. 		12	8	22	8		8			8		26e	
2. 	10a												13g
3. 													
4. 		20											
5. 	10b,19			8,23	20	20		8,22,23	8,19	22	8,19c,19d	8,26f	13h,13i, 13j,17
6. 						8							

*Note: The curves in this table illustrate the following relationships between volume and outcome:

1 Volume grouped into two categories, with lower rates of poor outcomes at higher volumes.
2 Downward-sloping line (negative linear relationship), with lower rates of poor outcomes at higher volumes.
3 U-shaped curve, with higher rates of poor outcomes at lower volumes, lower rates of poor outcomes in intermediate volume ranges, and higher rates of poor outcomes at higher volumes.
4 L-shaped curve (inverse logarithmic form), with large reductions in rates of poor outcomes as volume increases from low levels and little change at higher volumes.
5 Flat, indicating no significant relationship between volume and outcome; and
6 Upward-sloping line (positive linear relationship), with higher rates of poor outcomes at higher volumes

The numbers in the entries for this table refer to the numbered references in table 8-1 The meanings of the letters in the entries are as follows:

a Elective aneurysm surgery
b Ruptured aneurysm surgery
c Stomach operations, cancer diagnosis
d Stomach operations, ulcer diagnosis
e Postoperative deaths
f Major complications
g Femoral popliteal bypass, amputation outcome
h Carotid endarterectomy
i Aortofemoral bypass
j Femoral popliteal bypass, death outcome

SOURCE: Office of Technology Assessment, 1988

bines cholecystectomy with common bile duct exploration and other operations on the biliary tract. For CABG surgery, the hospital volume-mortality relationship may be driven primarily by emergency (study 25h) rather than scheduled patients (study 25k)(573). Similarly, Pilcher, et al., showed a volume-outcome relationship for ruptured (study 10b) but not elective aneurysm surgery (study 10d)(488).

Examining tables 8-2 and 8-3 overall makes it clear that a far greater number of available studies relate to hospital volume (table 8-2) than relate to physician volume (table 8-3); furthermore, many more studies of physician volume than of hospital volume found no relationship between volume and outcome. This pattern probably reflects three factors. First, physician volume data have been more difficult to obtain than hospital volume, so there have been more opportunities to undertake hospital studies. Second, even when data on physician volume have been available, it has been difficult to identify which physician is truly responsible for a patient when several specialists and consultants have been involved in a case. Third, some of the apparently inconsistent findings for physician volume may be due to methodological differences. Kelly and Hellinger (study 20), for example, found no surgeon volume-outcome relationship for cardiac catheterization when low-volume providers were omitted from their study. Hughes and his colleagues, focusing on low-volume surgeons, however, found worse outcomes associated with low-volume surgeons (study 8).

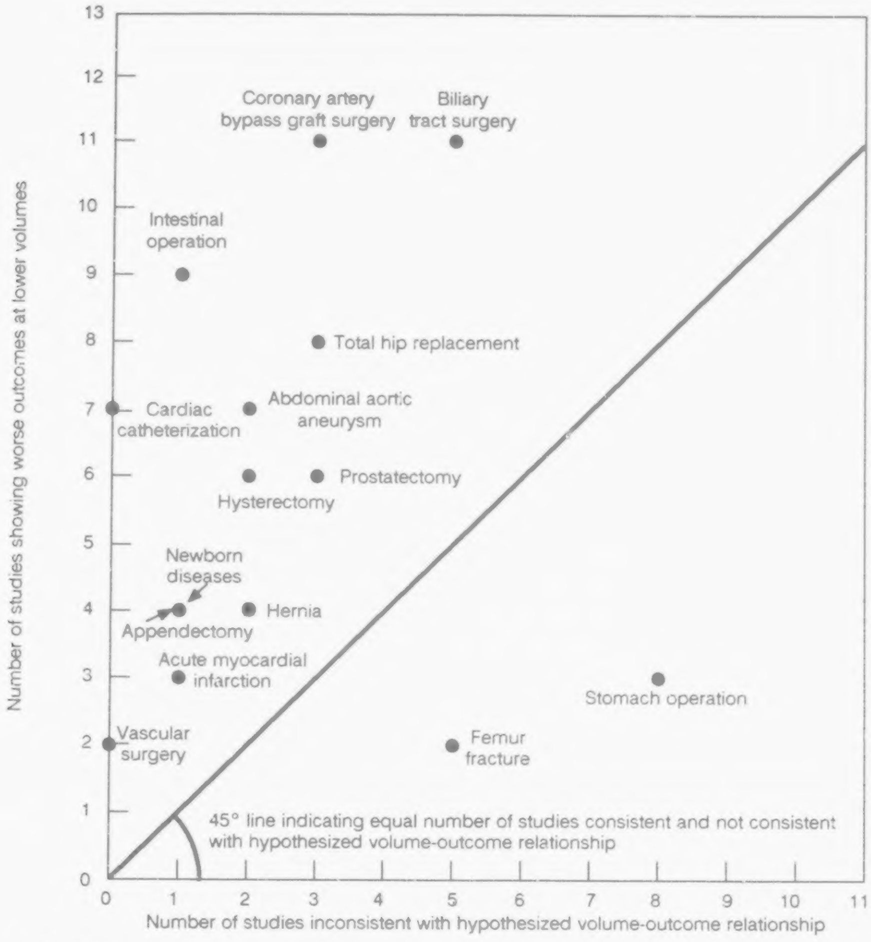
Without exception, none of the regression studies explicitly test a log versus a "U"-shaped curve, and there is little evidence of many observations on the upward sloping part of the "U." Therefore, it is possible to lump the first four types of findings as all supporting the notion that worse outcomes tend to occur in low-volume settings. (This is not necessarily the same as saying that more is better.)

Two types of results across procedures and diagnoses are summarized in figure 8-3: findings that are consistent with the hypothesis that worse outcomes occur at lower volumes and findings that are inconsistent with that hypothesis. For hospital volumes of abdominal aortic aneurysm, for example, there are seven studies indicating worse outcomes at lower volumes (Y axis) and two studies showing no relationship between volume and outcome (X axis). For each of the 13 diagnoses and procedures in the upper left half of figure 8-3, there are more studies showing worse outcomes at lower volumes than studies showing inconsistent findings with regard to the hypothesized volume-outcome relationship. Worse outcomes are demonstrated at lower volumes in 11 of 14 studies of CABG, in 9 of 10 studies of intestinal operations, in 8 of 11 studies of total hip replacement, and in all 7 studies of cardiac catheterization. Only for the two procedures in the lower right half of figure 8-3 (femur fracture and stomach operation) are there more findings of no effect of volume on outcome than of worse outcomes at lower volumes.

Although detailed analyses of the methods used by each study reviewed by OTA are necessary to understand why results differ for a single diagnosis or procedure, several important factors help explain inconsistencies across studies: 1) physician vs. hospital volume, 2) causal linkages from volume to outcome or outcome to volume, and 3) the problem of detecting an effect if the rate of poor outcomes is low and the sample size is small.

Relatively little work has been done to distinguish various causal linkages in the volume-outcome relationship. Hospitals with high volumes are often institutions in which physicians have high volumes, and it may be physician volume that truly matters. Therefore, it is crucial to distinguish between effects due to hospitals and effects due to physicians. Of the 124 findings concerning the effect of hospital volume on outcomes, 100 pertained to hospital volume without including physician volume, and 24 pertained to hospital volume and physician volume concurrently. Almost three-quarters of the 100 studies of hospital volume alone indicated a hospital effect, while only about half of the studies testing hospital and physician effects concurrently indicated a hospital effect. It appears, therefore, that in some instances, a measured hospital effect may be substituting for an untested physician effect. Alternatively, the high collinearity between physician and hospital volume may make it impossi-

Figure 8-3.—Number of Studies Reviewed by OTA Showing Either Worse Outcomes at Low Hospital Volume or No Effect, by Diagnosis or Procedure



SOURCE: Office of Technology Assessment, 1988

ble to detect true effects. Given the paucity of physician volume studies, one should reserve final judgment on this issue.

The uncertainty in this area reflects our limited understanding of the underlying reasons for the observed relationship between volume and outcome. The "practice-makes-perfect" explanation of the volume-outcome relationship rests on the general notion that increased experience results in more finely developed skills and, therefore, in better outcomes. The surgeon who consistently performs many units of a specific procedure will maintain, or continue to improve, his or her skills, while the surgeon who performs few procedures will become progressively less proficient. Similarly, nursing and other staff who are more familiar with certain types of patients may become or remain more proficient in working with them. Higher volumes may also make it possible for hospitals to purchase specialized equipment for such patients (217). Determining why outcomes for patients undergoing specific surgical procedures are related to volume requires extensive reviews of patients' medical charts from a large number of hospitals across a large number of procedures and diagnoses, because detailed data are unavailable from discharge data sets. For some procedures, problems in surgical technique may be the crucial factor, while for other procedures, inadequate postoperative monitoring may cause poor outcomes.

Even if physician volume is most important, hospital volume is likely to play a role. For example, a hospital with several high-volume and several low-volume surgeons may develop monitoring methods and standard procedures for the staff that catch errors and institute corrective actions. Thus, a low-volume surgeon may be "protected" in a high-volume hospital. Likewise, a surgeon with a high volume across several institutions but low volumes in each may achieve good results. The empirical testing of such hypothetical relationships is difficult because of the need to track data on the same physicians across hospitals.

Volume may not matter at all, but instead may serve as a marker for hospitals or physicians with special skills whose better-than-average performance attracts a disproportionate share of the referrals. This "selective-referral" hypothesis holds that any inverse relationship between volume and outcome arises from the attraction of more patients to physicians and hospitals with better outcomes. The idea that patients in some instances may look for hospitals or physicians with the best results seems implausible to some, who claim that the variation in mortality by disease or procedure is too small to influence patients' choice (218). If complications are correlated with mortality, however, variations in outcomes may be large enough to be noticed by patients' primary physicians who choose specialists for referral. Although it is difficult to identify an individual hospital or physician as having significantly worse than average death rates (396), referral patterns may be based on a simpler set of decision rules. If primary physicians switch referrals after even one "bad outcome," patients eventually are directed away from providers whose outcomes are worse than average.

Furthermore, even if the majority of patients go to the nearest hospital or otherwise make decisions independently of perceived outcomes, a minority seeking or referred to the "best provider in town" (or referred away from "poor-quality providers") will result in a selective referral pattern for specific diagnoses and procedures. As a result, hospitals with better outcomes would have higher-than-expected volumes. The question, therefore, is whether some patients are influenced in their choice of physicians and hospitals by relative performance, not whether all patients are so influenced.

Another principally empirical objection to the selective-referral hypothesis is that some studies show little relationship between outcomes and hospital characteristics traditionally considered to be markers of good performance, such as teaching status or board certification of physicians (217,393). However, these measures are rather blunt and unvalidated indicators of special expertise. It is common for a teaching hospital to be outstanding in the treatment of one diagnosis or procedure (e.g., cardiovascular surgery) but not to be particularly distinguished in another (e.g., neurosurgery).

When one attempts to test in a simultaneous-equation model both the effects of volume on outcomes and the effect of outcomes on volume, one may observe statistically significant effects for only one causal path. Even if the results indicate just an effect of outcome on volume in such a model, a simple test of volume as a function of outcome alone would probably show a relationship. There is not yet enough work to clearly indicate which causal paths are truly valid.

In designing an experiment, one should undertake a power test (ideally ahead of time) to determine the likelihood of detecting an effect if one truly exists. A power test is based upon the overall likelihood of the outcome's being measured and the sample size. There are substantial differences across studies in the number of patients involved and the average poor outcome (or mortality) rate.

To provide a sense of the issue at hand, consider the research findings from the 11 studies that reported on the hospital-volume/outcome relationship for the total hip replacement procedure. Eight studies showed a relationship between worse outcomes and low-volume hospitals, while three studies found no effect of volume on outcome (see table 8-2 and figure 8-3). The three studies that showed no effect had smaller sample sizes—under 1,500 patients in two studies and under 10,000 patients in the other study—than the sample sizes of from 13,700 to 33,000 patients in the eight studies that did find an effect. The three studies that had findings inconsistent with the hypothesized volume-outcome relationship probably had insufficient power to detect an effect unless it was very large. The mixed results for total hip replacement are not surprising given the design of the studies.

In summary, the available studies reviewed by OTA provide rather substantial evidence that worse outcomes occur at lower volumes for most of the procedures and diagnoses that have been studied. However, the volume-outcome relationship is not universal. For stomach operations and fractures of the femur, the evidence of a relationship is quite mixed, with the majority of studies indicating that volume has no effect on outcome. With the exception of the findings for stomach operations and femur fractures, all the other findings that suggest the lack of a relationship between volume and outcome either have low statistical power; are part of larger analyses in which a physician volume effect is found; or suggest a causal linkage from outcome to volume. Thus, although a relationship often exists, there is not yet enough evidence to distinguish effects due to physicians from effects due to hospitals or to have much confidence in the relative importance of the causal linkages.

FEASIBILITY OF USING THE INDICATOR

As has been discussed, there is frequently a relationship between volume and outcome. The general pattern is that better patient outcomes are associated with higher inhospital volumes. However, because there is hardly ever a perfect relationship, there are always some low-volume hospitals with apparently good outcomes and some high-volume ones with poor outcomes. This situation raises the obvious question, "How useful is volume as an indicator of the quality of care?" Since mortality data on Medicare patients are routinely available, why bother with volume data?

There will always be some chance component to a hospital's reported death rate in any single year, even after all adjustments for patient characteristics have been included. Various statistical calculations are designed to provide measures of this chance component and thus the degree of confidence one should have in the observed results for a particular hospital. It is inherent in the nature of small samples that one must expect much more variability in observed outcomes in hospitals with low volumes. One death among 10 or 20 patients may produce a mortality rate well above the average, but it is likely to be a chance occurrence. Similarly, even if the true or long-run mortality rate for that hospital is worse than average, with few patients in any particular year, there will often be years in which there are no deaths. To get a better estimate of the true performance of the outcomes in a low-volume hospital, one might aggregate data over several years, if they are available. Unfortunately, this technique makes it impossible to determine whether outcomes are improving or getting worse.

Combining data on volume and outcome is an alternative way of organizing a given amount of data to reduce the influence of chance and provide useful information. By aggregating data across hospitals within volume categories or using a regression to smooth out hospital-specific variability, the volume-outcome studies provide much more stable estimates of the performance of a class of hospitals. Although average results for all low-volume hospitals may not apply to a particular low-volume hospital, it is important to remember that, because of chance variability, last year's mortality rate for a particular hospital is not a very reliable indicator either. The two pieces of information, however, may be used together to guide a decision about a particular hospital.

The situation is different for high-volume hospitals, because the role of chance is smaller the larger the number of patients. Of course, hospital-specific mortality results will still be sensitive to unmeasured differences in patient characteristics that may not be adequately captured in the available data. If a high-volume hospital with worse-than-average outcomes claims that unmeasured patient-related factors account for the poor results, that claim may be worth more detailed investigation.

If high volumes for a particular procedure or diagnosis are primarily the result of superior outcomes, then the argument for volume data is even stronger. Since published hospital mortality data have only recently become available (see ch. 4), a relationship between volume and outcome implies that physicians (and possibly patients) have been able to use informal qualitative measures to guide more patients to physicians and hospitals with better results. Primary care physicians may consider both the mortality and other complications of their patients referred to certain specialists. Observations in the operating room or at the bedside may also alter one's confidence in the quality of care provided by specific physicians. Although such methods may be somewhat haphazard, they allow for a wide range of implicit but important criteria that may be valuable in the identification of which providers to seek out and which ones to avoid. It would be impossible to collect and make available such data, but if selective referral occurs, then the observation of a higher than expected volume of patients with diagnosis X in a hospital may be a valuable indicator of better-than-average quality.

It is important to note, however, that to use volume as an indicator of the quality of care, one must control for the various factors that influence volume. Large hospitals, for example, tend to have more patients of most diagnoses than small hospitals, irrespective of their relative quality. Public hospitals tend to treat a disproportionate share of diagnoses common among poor people. Selective contracts between certain payers and hospitals will also alter volumes. In much the same way that hospital-specific mortality rates are meaningless as outcome indicators until adjusted for case mix and certain other factors, hospital volumes are meaningless until adjusted for factors such as size of hospital, ownership, medical staff, and selective contracts. Although analyses with such adjustments have not yet been undertaken, they may be worth pursuing, especially for diagnoses and procedures for which there is evidence of selective referrals.

One additional use of volume as an indicator of the quality of care arises from the possibility of a volume-outcome relationship for physicians. Fewer studies have examined the volume-outcome relationship for physicians than have examined it for hospitals. Furthermore, the results for physicians are less consistent than those for hospitals, although some of the inconsistency may be due to methodological problems that can be overcome with better data and more analysis. Moreover, the problems of chance variation in small numbers of patients would make physician-specific data on mortality rates even less reliable than hospital-specific data. Volume data for physicians, however, may be far less controversial than outcome data. Thus, work on the volume-outcome relationship and familiarity with the use of hospital data could help set the stage for the use of physician volume data as an additional guide for consumers.

In choosing a physician or hospital, consumers should not just "go by the numbers." Instead, if there is good evidence of a volume-outcome relationship for the patient's specific diagnosis or prospective procedure, the patient should discuss the information with a primary care physician. Suppose, for example, that a physician is recommending that a patient have CABG surgery and there are several hospitals in the community with open-heart surgery teams. Even if hospital-specific mortality data are available, there may be questions as to how they should be interpreted if none of the hospitals have significantly high or low mortality rates. As proximity is not a major issue if there are several local hospitals and if the mortality rate (3 to 5 percent) is not trivial, the patient may want to find the best, or at least avoid the worst, institution.

Suppose the hospital initially selected had a low (but not significantly so) mortality rate last year, but this rate was based on only a small number of cases. If this hospital also had a low volume, it would be reasonable to press the physician on whether one of the higher volume centers with comparable mortality rates might not be more likely to have a lower true risk of a poor outcome. Such a question may encourage the physician to think further about the referral and perhaps informally seek out additional information about the best hospital to send the patient. Although this is a rather "soft" use of information, it is probably commensurate with the precision of the available data.

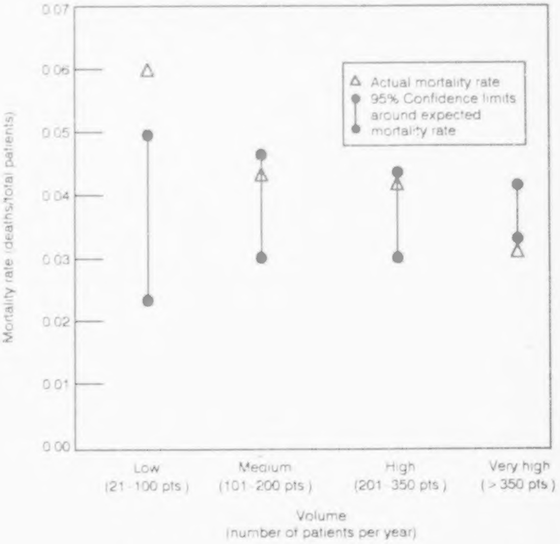
In using information about the relationship between volume and outcome, it is important to know the form of the curve for a particular diagnosis or procedure. In the analysis in this chapter, all findings with dichotomous results and with "downward-sloping," "L-shaped," and "U-shaped" curves were grouped together. If there truly is a "U-shaped" curve, then it is necessary to identify the volume level above which mortality rates begin to worsen. Several studies have estimated "U-shaped" curves, but none have tested whether a "U" was really superior to an "L" or similar form. Nor did the studies find much evidence that very high-volume hospitals actually had worse results. The only exceptions are the studies of outcomes for newborns by Rosenblatt, et al. (538) and Williams (702). In both instances, the authors argued that the apparently worse outcomes for newborns in the very high-volume hospitals could be attributed to the very high-risk infants referred to those hospitals pursuant to perinatal regionalization policies. Unless additional studies provide clear evidence that worse outcomes occur in very high-volume centers, the public need not worry too much about reports of "U-shaped" curves.

Even if outcomes do not get worse in very high-volume hospitals, available volume-outcome studies do not necessarily imply that more is better. In many instances, the rule might be: Avoid the very low-volume setting; once you find a hospital with a volume of X, there is little to be gained by looking for a hospital with higher volume. To make recommendations about specific optimal volumes would require analyzing up-to-date data on specific diagnoses and procedures across a wide range of hospitals. Unfortunately, the available published studies do not present such analyses, but the data are generally available and it would be relatively simple for an experienced research group to undertake the necessary analyses and make public the findings.

To provide a sense of how data might be presented, consider figure 8-4. (Similar data are published in a consumers' guide in the Washington, DC area (693).) The figure indicates age- and sex-adjusted mortality rates for patients undergoing CABG surgery in hospitals with various volumes and also shows the confidence intervals, the ranges in which mortality rates would be expected to fall if volume were not a factor. (Although adjusting for risk factors other than age and sex would improve the quality of the data, the presentation could be similar.) Mortality rates in the very highest volume hospitals are significantly lower than expected; part of the reason is that at higher volumes, the confidence interval narrows. Because hospital-specific mortality data are more reliable at high volumes, however, the volume data for hospitals with high volume are less valuable. Also, patients will be less willing to switch hospitals for the relatively small incremental improvement in expected mortality associated with very high-volume, in contrast to medium- or high-volume, hospitals.

Figure 8-4 also shows that patients undergoing CABG surgery in low-volume hospitals experience significantly higher than expected mortality rates. The difference not only is statistically significant, but it amounts to a half-again higher

Figure 8-4.—Comparison of Actual and Expected Mortality Rates for Patients Undergoing Coronary Artery Bypass Graft Surgery in California, 1983



SOURCE: J.A. Showstack, K.E. Rosenfeld, D.W. Garnick, et al., Institute for Policy Studies, University of California, unpublished data, San Francisco, CA, 1987 rate—a 6-percent mortality rate instead of a 4-percent rate. More importantly, because of the problems of chance variability in mortality rates, review of hospital-specific mortality rates would identify few of the low-volume hospitals as having significantly poor hospital-specific outcomes. Thus, both hospital-specific mortality data and more general volume-outcome information are helpful in guiding consumers to ask better questions of their physicians.

The use of volume and outcome data varies with the specific situation at hand. In many situations, hospitalization and treatment must be immediate, and there is little time for discussion, let alone referral of a patient to other settings. In other situations, however, there may be time for reflection and discussion, but the evidence may suggest only a very weak relationship between volume and outcome. Although this relationship may be statistically significant because of the large data sets used for the analysis, the difference between an average mortality rate of 1.0 percent and 1.1 percent may not be worth pursuing for some patients, especially since there may be other factors of importance, such as proximity, the retention of a well-trusted family physician, or an institution's reputation for having attentive and responsive nursing staff.

CONCLUSIONS AND POLICY IMPLICATIONS

OTA's review of the research literature on the volume-outcome relationship for hospitals and physicians suggests that, at least for some diagnoses and procedures, higher volumes are associated with better outcomes. For 13 procedures and diagnoses reviewed in OTA's literature survey, more than half of the studies focusing on hospital volume showed this relationship. For only two procedures, femur fractures and stomach operations, did a majority of studies show no relationship between volume and outcome. The evidence for hospitals overwhelmingly showed worse outcomes at lower volumes for CABG surgery, intestinal operations, total hip replacement, cardiac catheterization, abdominal aortic aneurysm, and biliary tract surgery. Fewer studies focused on physician volume than on hospital volume, and more of the studies on physician volume either had inconsistent findings or showed no effect of volume on outcome.

To some extent, it is difficult to determine whether volume is a useful indicator of the quality of care because of the continuing controversy over the relative importance of 1) increased volume's providing the opportunity for practice and thus better outcomes, and 2) intrinsically better providers' generating increased volume through referrals. The repeated observation of a simple association between volume and outcome does not help distinguish between these two hypotheses or reveal any other causal mechanisms.

Photo credit: Cleveland Clinic Foundation

Lower volume of coronary artery bypass graft surgery in hospitals was associated with higher mortality rates in 11 of 14 studies reviewed by OTA.

Regardless of the true causal pathway, volume information is useful as an indicator of quality. If the only influence is of practice, consumers would usefully be directed toward more experienced practitioners. If the influence is primarily from good outcomes' generating higher volumes, then volume may be even more valuable to consumers, because high volumes generated by selective referrals may be the best indicator of good quality.

Research focused in specific areas could provide necessary further information about the volume-outcome relationship. A problem with much of the research to date has been its academic focus; investigators have explored various analytic questions rather than developing sets of estimates that are directly useful for consumers. For example, although many studies indicate the presence of a volume-outcome relationship, the wide range of analytic methods and variable specifications makes it difficult to determine whether poor outcomes are concentrated at very low volumes or whether improved outcomes are seen throughout the observed range of volumes. Further studies are required to determine whether the recommendations should be to "seek the highest volume center" or to "avoid places with fewer than X patients."

To some extent, the variety of functional forms and approaches used by various investigators reflects the constraints of the available data. A very useful study would compare the findings of studies that used the same analytic techniques on various types of data for patients who had the same diagnosis or procedure. For example, one can obtain data on post-discharge mortality and readmissions for Medicare patients, but using these data limits the analysis to patients over age 65. Are linked inpatient and ambulatory data superior to data on inpatient outcomes? In a similar vein, do hospitals with high rates of other complications also have high mortality rates? Do these objective measures match other evaluations of quality, such as those developed by the peer review organizations?

The quality of the data is probably more important for the evaluation of specific hospitals than for the analysis of volume-outcome issues. Additional data that may improve the certainty of a judgment with respect to quality of a particular hospital are very important because of the potentially disastrous consequences of misclassification. In contrast, random noise in the data used for volume-outcome studies merely makes it somewhat more difficult to detect what is going on; a larger sample size can often overcome the problem.

The evidence of a relationship between physician volumes and outcomes is less clear than the evidence for hospital volumes and outcomes, and none of the existing studies of physician volumes is fully convincing. Prior research has been constrained by both data and methodological problems. Some newly available data sets are now including physicians' license numbers (and in Arizona's case, physicians' names), so it will be possible to identify a physician's patients across several hospitals.

Another crucial question that remains to be resolved is whether high volumes arise from selective referrals of patients to hospitals and physicians with better-than-average outcomes, whether better outcomes arise from high volumes, or whether both phenomena arise in some complex relationship. Methodologically, this question is a difficult one to address, but various simultaneous-equation techniques and better data on patient referrals may provide more convincing evidence.

Both selective-referral and practice-makes-perfect effects have a time dimension. The fact that a beginning surgeon will eventually perform 200 procedures during the first year of practice may not affect the outcomes of the first patients on which he or she operates in that year. It is often assumed, however, that a hospital or physician with a volume of 200 procedures in a year had about that many procedures in prior years. The implicit assumption is that all volumes have reached some steady-state level. In reality, new physicians enter practice, new procedures are developed, hospitals offer new services, and past volume levels may differ from current (and future) ones. What is the shape of the personal and institutional learning curves after a new procedure or treatment is introduced? What yearly volume is necessary to keep skills from deteriorating?

For considering the selective referral hypothesis, timing is also important. If outcomes or reputations influence referrals, what is the time lag involved? Is an occasionally higher-than-expected mortality rate ignored, while only consistently better or worse than average results affect referrals? Can a hospital that replaces a poor-quality surgeon with a good-quality one increase its volume, or is a poor reputation difficult to erase? Likewise, for how long can a hospital (or physician) with deteriorating outcomes maintain old referral sources? These questions have not been explored in any empirical studies to date.

Finally, a series of very detailed studies could explore precisely what clinical factors account for differences in outcomes, in effect, to validate the observed relationship between volume and outcome. Such studies would probably rely on careful review of patients' charts from various settings to determine the relative importance of errors of commission and omission, differences in technique, monitoring, support, and the like. It is probable that the importance of various factors will depend on the procedure or diagnosis studied.

Even with the substantial gaps in knowledge about the volume-outcome relationship, there are still policy measures worthy of consideration. In discussing various policy options, it is important to consider unintended incentive effects. The following five policy measures are ordered roughly in terms of increasing strength of incentives for—and ability of—hospitals to manipulate the data or otherwise behave in undesirable ways.

Educating the general public about the relationship of lower hospital volumes to worse outcomes is the simplest approach. Even if the causal linkages are not clear, it seems reasonable to argue that, in the absence of other evidence, hospitals with high volumes are preferable to nearby ones with very low volumes. Upon receiving a referral for a specialized procedure, an informed consumer might then ask his or her primary care physician about the volume and quality of the proposed specialist and hospital, given the relevant alternatives. Educating the general public would impose no new data collection requirements and the potential costs are small. One could easily see an educational strategy implemented through articles in the lay press, such as *Readers' Digest* or the Sunday newspaper supplement.[5]

A second level of intervention might be directed toward physicians through their specialty associations and continuing-education programs.[6] Specialty associations might be encouraged by Congress to collect volume and outcome information in their areas and make it available to physicians. In particular, these associations could focus on some of the more individualistic and sensitive factors that may improve physicians' ability to selectively refer patients to settings and physicians with better outcomes. It might be necessary to clarify whether such educational efforts by local specialty associations would raise antitrust problems.

A third level of intervention would be for States or other State-level entities to require the routine collection and publication of hospital-specific volume information. For the 28 States with mandatory hospital discharge abstract reporting requirements, this task would be an easy one. States could clearly not publish data for all hospitals and all procedures and diagnoses, but selected data could be made available to interested parties. Selected hospital-specific information could be reprinted by local newspapers. California Blue Shield published a list of hospitals with their CABG surgery volumes (112), and Blue Cross and Blue Shield of Ohio published a Consumer Guide with the number of patients by DRG (73). Some consumer organizations and magazines have done the same (693).

Requiring the disclosure of hospital-specific volumes is a measure that must be carefully considered because of potential unanticipated effects. Consider, as an example, a fourth level of intervention whereby a hospital is penalized financially by third-party payers, or a particular unit shut down by regulators, if certain volume levels are not maintained. This approach would create incentives for hospital administrators to make sure that at least the minimum acceptable number of patients are treated. One could imagine memos from hospital administrators to the medical staff pointing out that if another 20 patients are not operated on before the end of the fiscal year, the X unit will be closed down. This pressure might lead to the relaxation of standards for the appropriateness of an admission. Moreover, basing payment or regulatory decisions, which affect a hospital's ability to continue in a specific line of business, on volume may not be fair because volume is at best merely a proxy for quality.

---

[5] Information-dissemination strategies are discussed in ch. 2.

[6] Physician specialty boards are listed in ch. 10.

A fifth policy application using volume data as an indicator of the quality of care is in the realm of selective contracting. Insurers, health maintenance organizations, and other agents such as Medicaid programs may wish to steer the patients for whom they are responsible to hospitals that are likely to achieve better outcomes. If reliable outcome data are available either through sources that routinely collect data or through carefully structured bids, then for high-volume hospitals, outcome data may be preferred to simple volume data because the outcome data would include only a small chance component. For low-volume hospitals, the outcome data tend to be too unreliable. On the other hand, if outcome data are unavailable or too subject to manipulation, then volume of specific procedures may be a proxy for quality. (For example, suppose an agency were to announce that it was going to utilize hospital discharge abstracts to determine death rates for the purposes of contracting. A hospital with a high inpatient mortality rate may monitor patients for complications and transfer those at risk of death, thereby improving its own statistics. It would be far more difficult to manipulate volume figures, and it is unlikely that many hospitals could attempt such a strategy without detection.)

Additional policy applications depend on a better understanding of the relationship between volume and outcome. For example, if increasing volume for specific procedures or diagnoses does lead to improved outcomes, then the argument for explicit regionalization strategies becomes far stronger. If hospital volume is far more important than physician volume, then one would argue against the peripatetic surgeon. On the other hand, if physician volume is the crucial variable, then "circuit riding" may become far more common, with many low-volume hospitals sharing a single high-volume physician. If higher hospital malpractice claims are associated with lower volumes, then malpractice insurance premiums should be adjusted to reflect this risk factor. All of these and other options must await future research. Fortunately, many of the policies directed toward consumers do not require much additional information.

Chapter 9

Scope of Hospital Services: External Standards and Guidelines

CONTENTS

Page
Introduction 187
External Standards and Guidelines 187
  Standards for Overall Hospital Accreditation/Certification 187
  Standards and Guidelines for Specific Services 190
Reliability of the Indicator 197
Validity of the Indicator 197
  Validity of Overall Hospital Accreditation 199
  Validity of Standards and Guidelines for Specific Services 200
Feasibility of Using the Indicator 200
Conclusions and Policy Implications 202

Box

Box *Page*
9-A. Selected Sources of Information About Scope of Hospital Services 203

Tables

Table *Page*
9-1. Various Organizations' Standards and Guidelines for Staffing Neonatal Care Facilities 192
9-2. Characteristics of Various Organizations' Standards and Guidelines for Emergency Services 194
9-3. Specialty Organizations To Be Consulted in Developing the American Medical Association's "Guidelines for Classification of Hospital Emergency Capabilities," January 1988 195
9-4. HCFA's Condition of Participation Governing Emergency Services 195
9-5. Characteristics of Trauma Center Designations by State 198
9-6. States That Require Copies of JCAHO Accreditation Reports From Hospitals 201
9-7. Characteristics of External Standards and Guidelines for Hospitals: Overall Accreditation/Certification and Specific Services 204

Chapter 9

Scope of Hospital Services: External Standards and Guidelines

INTRODUCTION

Scope of hospital services is a structural measure that reflects whether a hospital has the resources—facilities, staff, and equipment—to provide care for the medical conditions it professes to treat or to care for the medical conditions affecting potential patients. There are several potential sources of information on the scope of a hospital's services, including hospital advertising, media reports about the existence of special equipment or specially trained staff, consumer guidelines for selecting medical providers, and organizations that accredit or certify hospitals.[1] Identifying whether a hospital complies with external standards such as those used for accreditation or certification by an external body, however, is likely to be the most valid means of ascertaining a hospital's scope of services. Accreditations and certifications for scope of hospital services are distinct from some of the other indicators evaluated in this report in at least one sense. As currently constructed, they measure only the capability of a hospital to deliver good quality care, not the quality of care actually delivered or its outcome.

This chapter briefly describes two national methods of overall accreditation/certification of hospitals, that of the Joint Commission on the Accreditation of Healthcare Organizations (JCAHO) and that of the Health Care Financing Administration (HCFA). It then describes external standards and guidelines for neonatal intensive care units, cancer care, and hospital-based emergency and trauma services. The next sections of the chapter analyze the reliability, validity, and feasibility of using external standards and guidelines related to the scope of hospital services as indicators of the potential of a hospital to deliver good quality care. The final section draws conclusions and discusses policy implications.

EXTERNAL STANDARDS AND GUIDELINES

Standards for Overall Hospital Accreditation/Certification

JCAHO Accreditation

The most well-known and widely applied hospital accreditation standards are those of JCAHO. Of the approximately 6,800 hospitals of all types in the United States, about 5,000 (70 percent) are surveyed by JCAHO. Submitting to JCAHO evaluation is voluntary, but not all hospitals are eligible for JCAHO surveys (325). One reason that JCAHO accreditation is important is that such accreditation, along with certain additional criteria, is a condition of participation in the Medicare and Medicaid programs (Section 1865 of the Social Security Act).[2] Medicare and Medicaid pay for

---

[1] Hospital certification typically refers to approval by governmental bodies; accreditation usually indicates approval by a private organization, most often a professional organization of peers. The term "guidelines" refers to standards proposed by professional organizations and voluntarily applied by providers.

[2] In addition to being accredited by JCAHO, hospitals must meet requirements for utilization review (Section 1861(e)(6) of the Social Security Act (42 CFR Subpart S, 405.1901(d)(1) and 482.30)) and discharge planning (Public Law 99-190). In practice, the require-

(continued on next page)

about 38 percent of the hospital care provided in this country (715). JCAHO accreditation is also woven through the hospital licensure requirements of 41 States (323) and is a condition of participation for an unknown number of insurance companies (48).

JCAHO conducts a complete survey of each eligible hospital once every 3 years and assesses each hospital's compliance with over 2,000 standards. The purpose of the JCAHO hospital accreditation process is to evaluate each hospital's *overall* capability of providing medical care. Thus, particular attention is paid to functions affecting the entire hospital, such as the governing body, the medical staff, nursing services, infection control, and quality assurance, and the way these and other functions are integrated across the hospital. Throughout this chapter, and for purposes of evaluating JCAHO accreditation as a potential indicator of the quality of care, it is important to keep in mind that it is not JCAHO's purpose to separately accredit individual hospital departments such as those that provide emergency services or neonatal intensive care. Because JCAHO does survey and evaluate those services as part of its overall accreditation process, however, JCAHO standards for these separate departments are discussed in this chapter as having the potential to evaluate whether hospital scope of services is appropriate.

JCAHO standards are developed by panels of experts, sometimes with the aid of scientific literature, and are evaluated by interested hospitals and other experts before their adoption. JCAHO standards and required characteristics focus on certain key functions across the hospital: quality assurance, privilege delineation, existence of policies and procedures, and infection control.

A hospital's failure to comply with key JCAHO standards sometimes results in "accreditation with contingencies." JCAHO gives each hospital a contingency score (which may be zero) that determines in part whether the hospital is accredited. The actual accreditation decision is made by JCAHO's Accreditation Committee, following a recommendation by JCAHO staff, using a set of weighting procedures and objective rules to ensure consistency across hospitals. If a hospital receives a contingency, it must satisfy JCAHO within a specified period of time that it is in compliance with the problem standards. Depending on the nature of the contingencies, hospitals may have to submit to a focused resurvey, usually within 6 to 9 months from the date they receive the report. From 1982, when the current JCAHO accreditation procedure was implemented, until 1987, the percentage of surveyed hospitals with JCAHO contingencies of any type increased from about 65 percent to 90 percent (387,388,524).

About 75 hospitals (5 percent of JCAHO-surveyed hospitals) each year receive enough contingencies of a serious nature that a formal nonaccreditation decision from JCAHO looks probable; the hospitals are informed of this possibility by JCAHO staff before the staff recommendation goes to the JCAHO Accreditation Committee. Among the 5 percent, 3 to 4 percent of the JCAHO-surveyed hospitals correct their deficiencies to the satisfaction of JCAHO and avoid a formal nonaccreditation decision. Each year, about 1 to 2 percent of all JCAHO hospitals surveyed, or 15 to 30 hospitals, are formally judged by JCAHO to be nonaccredited. Some of the 1 to 2 percent of hospitals that are formally nonaccredited work on correcting deficiencies while they are appealing the JCAHO decision and then request a resurvey; others drop their quest for JCAHO accreditation, sometimes permanently. Some hospitals do, however, request HCFA inspection following nonaccreditation by JCAHO.

HCFA Certification

Hospitals that desire Medicare and Medicaid reimbursement but choose not to be surveyed by

---

(continued from previous page)

ments for utilization review are met by the existence of utilization and quality control peer review organizations.

In general, to meet the Medicare and Medicaid conditions of participation, hospitals must meet "any requirement under section 1861(e) of the [Social Security] Act and implementing regulations which the Secretary [of Health and Human Services], after consulting with [the Joint Commission] and [the American Osteopathic Association], identifies as being higher or more precise than the requirements for accreditation (section 1865(a)(4) of the Act)" (42 CFR Subpart S, 405.1901(d)(3)). Psychiatric hospitals must meet "the additional special staffing requirements that are considered necessary for the provision of active treatment in psychiatric hospitals (section 1861(f) of the Act) and implementing regulations" (42 CFR Subpart S, 405.1901(d)(2)).

Photo credit: Joint Commission on the Accreditation of Healthcare Organizations

A JCAHO surveyor examines hospital records. JCAHO's accreditation process is intended to evaluate the overall capability of a hospital to provide medical care, rather than to evaluate particular services.

JCAHO *or* cannot meet JCAHO's eligibility or accreditation criteria may opt to be certified by HCFA. About 1,400 hospitals per year routinely choose to be surveyed by HCFA. Because for every day that a hospital is not certified by HCFA, it loses Medicare and Medicaid reimbursement, not being accredited by JCAHO or certified by HCFA is very costly for a hospital.[3]

Most HCFA-certified hospitals are small, rural community hospitals (438). Texas has the largest number of HCFA-certified hospitals (157 hospitals), followed by Kansas (83), Minnesota (63), Georgia (59), Nebraska (56), Mississippi (55), California (53), Oklahoma (51), Louisiana (50), Florida (48), and Iowa (46) (438). Those 11 States have half the non-JCAHO-accredited, HCFA-certified hospitals in the United States and its possessions.

HCFA uses survey methods that are somewhat different from JCAHO's. HCFA's hospital surveys are conducted annually, whereas JCAHO's are conducted every 3 years. HCFA/State surveyors have the force of law and the threat of noncertification to ensure compliance, while the JCAHO organization does not. HCFA surveyors are State personnel, and although the teams receive some training from HCFA, their composition is determined by the States (399). JCAHO surveyors are hired and trained by JCAHO. JCAHO provides 2 weeks of didactic training, a 3- to 4-week preceptorship, and an annual 3-day conference for surveyors.

JCAHO has stricter criteria for surveyors than does HCFA. JCAHO requires each survey team to include one physician, one nurse, and one hospital administrator. In addition, JCAHO requires the nurse and hospital administrator surveyors to have had administrative experience in a hospital. The qualifications of HCFA/State surveyors are more diverse, and many of these surveyors are not as highly trained as JCAHO surveyors. Of the 2,786 surveyors (of a total of about 3,400) who responded to a HCFA questionnaire, for example, only 10 (less than one-half of 1 percent) were medical doctors (646). Finally, HCFA has substantially fewer standards than does JCAHO, and HCFA's conditions of participation are much less detailed than JCAHO's standards. Generally, 1 percent or less of the hospitals surveyed by HCFA each year are terminated from the program involuntarily (249).[4]

---

[3] Accreditation by the American Osteopathic Association enjoys the same status with respect to Medicare and Medicaid payment as JCAHO accreditation.

Overall Hospital Accreditation/Certification and Scope of Services

Neither JCAHO accreditation nor HCFA certification is designed to assess whether particular hospital departments are capable of providing specific services. Nevertheless, JCAHO accreditation or HCFA certification does ensure that a certain scope of services exists in a hospital. In order to qualify for the survey on which JCAHO accreditation is based, a hospital must meet certain eligibility criteria. The hospital must maintain facilities, beds, and services that are available over a continuous 24-hour period, 7 days a week. Unless a hospital is a psychiatric or substance abuse facility, it must also provide diagnostic radiology, dietetic, emergency, rehabilitation, and respiratory care services, among others. In addition, it must provide at least one of the following acute-care clinical services: medical, obstetric-gynecological, pediatric, surgical, psychiatric, or alcohol- or drug-abuse services. If the hospital provides obstetric-gynecological or surgical services, it must also provide anesthesia services.

A hospital is also required to supply far fewer hospital services for HCFA certification than for JCAHO accreditation. Services required by JCAHO that are not required by HCFA include emergency services, nuclear medicine services, some type of special care services, professional library services, and social work services. For both JCAHO accreditation and HCFA certification, surgical services are optional.[5] Although both HCFA and JCAHO rate a number of specific departments or services (e.g., diagnostic radiologic services, outpatient services, surgical and anesthesia services), for the most part, neither rates condition-specific services such as heart disease or cancer services.

Standards and Guidelines for Specific Services

Neonatal Intensive Care Services

In 1976, in the face of a proliferation of neonatal intensive care units, the Committee on Perinatal Health[6] proposed guidelines for the regionalization of U.S. maternal and perinatal health services (142). Underlying the concept of regionalization of these services is the idea that high-risk mothers and infants will be screened and referred or transported to the appropriate level of care. The Committee on Perinatal Health proposed three levels of hospital care for perinatal services. Hospitals that served as regional centers and provided the most sophisticated neonatal intensive care were to be designated Level III facilities. Hospitals that provided neonatal intensive care but lacked some services provided in Level III facilities were to be called Level II facilities; and hospitals that provided normal newborn care with no special units for the care of seriously ill infants were to be called Level I facilities.

In 1983, the American Academy of Pediatrics and the American College of Obstetricians and Gynecologists more fully explicated the responsibilities and requirements of the three levels of hospitals in the regional system of maternal and perinatal services. A document issued by these organizations specified guidelines for minimum number of beds, square footage per bed, personnel, hospital structure, equipment, ancillary support, and educational services for parents (15).

A recent analysis by OTA concluded that neonatal intensive care has been in large part responsible for the remarkable decline in U.S. neonatal mortality rates over the past 25 years and has contributed to improved long-term developmental outcomes for premature infants; the improved survival of premature infants has not been accompanied by an increase in the proportion of babies with serious long-term disability (194).[7] According to OTA's analysis, however, "an extremely premature baby's chances for survival and normal development are in large part determined by where the baby is born" (194). The evidence strongly suggests that the likelihood of survival among very low birthweight babies (babies weighing under 1,500 grams at birth) is highest if the baby is born in a hospital designated a Level III neonatal facility. When considering these conclusions, however, one should keep in mind that they are based on some studies that were not methodologically rigorous (i.e., studies that did not use random assignment of newborns to compare Level I, II, or III facilities). Some studies have found that very low birthweight infants in Level III units had lower mortality rates than those in Level II units.

---

[4] In fiscal year 1987, 9 hospitals were terminated involuntarily, in fiscal year 1986, 20 hospitals were terminated involuntarily, and in fiscal year 1985, 8 hospitals were terminated involuntarily for not meeting HCFA's conditions of participation.

[5] The reason surgical services are optional for JCAHO is to make it possible for psychiatric hospitals to be accredited. In most other respects, psychiatric hospitals are held to the same standards as all other accredited hospitals.

[6] The Committee on Perinatal Health was a joint effort by the American Medical Association, the American College of Obstetricians and Gynecologists, the American Academy of Family Physicians, and the American Academy of Pediatrics.

The concept of regionalization for perinatal services has not been so well accepted by hospitals and physicians, however (194). Despite the existing guidelines, there is no standard national application of what constitutes Level II or Level III perinatal care (106). Ohio and some other States use the American Academy of Pediatrics/ American College of Obstetricians and Gynecologists guidelines to evaluate each hospital's perinatal services and assign levels accordingly (73,106). In California and most other States, however, the regional system of perinatal services is informal, and each hospital classifies its own services (344).

JCAHO applies standards for neonatal intensive care units in its overall hospital accreditation process (325), but these JCAHO standards are much less detailed and specific than the guidelines of the American Academy of Pediatrics and American College of Obstetricians and Gynecologists. Table 9-1 illustrates some of the differences between them in terms of staffing. JCAHO's standards do not differentiate between Level II and III neonatal intensive care. Even though JCAHO evaluates neonatal or other specific services as part of its overall hospital accreditation process, consumers may want to go beyond JCAHO accreditation to approvals by the appropriate specialty organization.

Cancer Care

Being stricken with cancer creates great fear among patients, and patients with cancer are intensely interested in finding the appropriate place for treatment. At least three organizations of independent observers have devised systems of approval for cancer treatment centers:

- the American College of Surgeons,
- the Association of Community Cancer Centers, and
- the National Cancer Institute.

There are substantial differences among them.

Cancer program approval by the American College of Surgeons is granted following an application and a survey by three members of the Commission on Cancer. The four basic requirements for American College of Surgeons approval are as follows:

1. the existence of an established multidisciplinary cancer committee that meets quarterly and provides the overall leadership of the cancer program;
2. an established tumor registry with 2 years of patient data and 1 year of successful (minimum 90 percent) patient followup;
3. patient-oriented, multidisciplinary cancer conferences conducted weekly or monthly; and
4. completion of two patient care evaluation studies each year (27,28).

Failure to comply with any one of these requirements results in either a 1-year approval (versus the usual 3-year approval) or, if there are other significant deficiencies, nonapproval. When a hospital first applies for approval, approval is not granted if there are any deficiencies. The American College of Surgeons has approved about 1,200 cancer programs (356), and an additional 400 to 500 are awaiting approval (469a). Centers approved by the American College of Surgeons'

---

[7] OTA did find, however, that there has been an increase in the absolute number of survivors with serious long-term disability (194).

Table 9-1.—Various Organizations' Standards and Guidelines for Staffing Neonatal Care Facilities

JCAHO Standards for Staffing Neonatal Intensive Care Units

S.P.7.4.2. The director or other qualified physician designee in charge of the unit has at least 1 year of recognized special training and experience, as well as demonstrated competence, in neonatology.

S.P.7.4.3. Pediatric surgery is provided in the hospital, as required.

S.P.7.4.4. Nursing care is supervised by a registered nurse who has training, experience, and documented current competence in the nursing care of high-risk infants.

S.P.7.4.5. The nursing staff is proficient in teaching parents how to care for their infants at home.

S.P.7.4.9. Radiologic technologists are familiar with X-ray techniques to be used with newborn infants so that repetitive exposures are not necessary.

American Academy of Pediatrics and American College of Obstetricians and Gynecologists' Guidelines for Staffing Level I, Level II, and Level III Neonatal Facilities

Level I	*Level II*	*Level III*
Chief of service One physician responsible for perinatal care (or codirectors from obstetrics and pediatrics)	Personnel Joint Planning: Ob: Board-certified obstetrician with certification, special interest, experience, or training in maternal-fetal medicine; Peds: Board-certified pediatrician with certification, special interest, experience or training in neonatalogy	Codirectors: Ob: Full-time board-certified obstetrician with special competence in maternal-fetal medicine. Peds: Full-time board-certified pediatrician with special competence in neonatal medicine
Other physicians: Physician (or certified nurse-midwife) at all deliveries Anesthesia services Physician care for neonates	Level I plus: Board-certified director of anesthesia services Medical, surgical, radiology, pathology consultation	Levels I and II plus: Anesthesiologists with special training or experience in perinatal and pediatric anesthesia Obstetric and pediatric subspecialists
Supervisory nurse Registered nurse in charge of perinatal facilities	Ob: RN with education and experience in normal and high-risk pregnancy only responsible Peds: RN with education and experience in treatment of sick neonates only responsible	Supervisor of perinatal services with advanced skills Separate head nurses for maternal, fetal, and neonatal services
Staff nurse/patient ratio Normal labor 1:2 Delivery in second stage 1:1 Oxytocin inductions 1:2 Cesarean delivery 2:1 Normal delivery 1:6-8	Level I plus: Complicated labor/delivery 1:1 Intermediate nursery 1:3-4	Levels I and II plus: Intensive neonatal care 1:1-2 Critical care of unstable neonate 2:1
Other personnel Licensed practical nurse, assistants under direction of head nurse	Level I plus: Social service, biomedical, respiratory therapy, laboratory as needed	Level I plus: Designated and often full-time social service, respiratory therapy, biomedical engineering, laboratory technician Nurse clinician and specialists Nurse program and education coordinators

SOURCES: **JCAHO standards:** Joint Commission on the Accreditation of Healthcare Organizations, *AmH/88: Accreditation Manual for Hospitals* (Chicago, IL: 1988); **AAP/ACOG guidelines:** American Academy of Pediatrics and American College of Obstetricians and Gynecologists, *Guidelines for Perinatal Care* (Evanston, IL: 1983).

Commission on Cancer are listed in the American Hospital Association's *Guide to the Health Care Field* (29), and a list of approved programs is available from the American College of Surgeons. The American College of Surgeons does not disclose how many programs have been refused approval, except to say that the number is small.

The American College of Surgeons' patient care evaluation studies are similar to JCAHO's monitoring and evaluation requirements,[8] except that established American College of Surgeons-approved programs are required to complete one study to measure process and one study to measure outcome, and new programs may complete two process studies each year until sufficient data are available to participate in the outcome studies. The specifics of both JCAHO's and the American College of Surgeons' monitoring/evaluation programs are determined internally at the hospital.[9] Unlike JCAHO, however, the American College of Surgeons requires that the outcome study compare the hospital's experience with national or regional results (27). The American College of Surgeons does have a voluntary program of cancer patient care evaluation, in which results are compared across hospitals.

---

[8] A key aspect of hospital quality assurance activities required by JCAHO, the monitoring and evaluation process includes identifying important aspects of care, identifying indicators related to these

In comparison to the American College of Surgeons' program, the accreditation program of the Association of Community Cancer Centers is just beginning. Membership in the Association of Community Cancer Centers is granted if a cancer center has the following:

1. a multidisciplinary cancer program;
2. supervision by a multidisciplinary cancer committee, group, or team; and
3. direct or indirect involvement with care for cancer patients.

Membership is open to freestanding cancer centers, health maintenance organizations, physician group practices, home health agencies, hospital-based cancer programs and individual providers (45). The Association of Community Cancer Centers has standards, but they operate primarily as guidelines to be used as self-assessment tools by the association's organizational members (46). The Association of Community Cancer Centers has about 30 hospital members and plans to begin a survey process in the near future (179).

The National Cancer Institute has several programs to designate cancer centers: the Comprehensive Cancer Centers program, the Community Clinical Oncology Program, and the Cooperative Group Outreach Program. Such designations are a requirement for receiving support grants and are based primarily on research capability (118, 580, 668).

Emergency and Trauma Services

Emergency and trauma services involve situations in which life or death may be at stake, and are therefore of extreme importance to consumers. In addition, consumers seem more likely to choose an emergency department than other hospital departments, although they may consult their physicians for advice or direction.[10]

There are several sources of standards and guidelines for the scope of emergency services that may potentially be of use to consumers. JCAHO, HCFA, the American College of Emergency Physicians and Emergency Nurses Association (ACEP/ENA), and the American Medical Association (AMA) all have or are planning guidelines for emergency services (23,36,38,325,642).[11] The American College of Surgeons has a set of guidelines for trauma care (26). In addition, many States and other localities have requirements that hospitals must meet to provide emergency services and/or to be designated as trauma centers.

Here as in other sections of this chapter, the distinction must be kept in mind between standards and guidelines. Only the requirements for emergency services and trauma care of JCAHO, HCFA, and States and localities are required for accreditation or certification by those organizations and can strictly be considered standards. Specialty organizations provide guidelines for emergency services and trauma centers, but their use by hospitals is optional. The ACEP/ENA guidelines, for example, are a "statement of suggested capability . . . not designed to be interpreted as mandatory by legislative, judicial, or

---

aspects of care, establishing thresholds for evaluation related to the indicators, collecting and organizing data, evaluating care when thresholds are reached, taking actions to improve care, assessing the effectiveness of the actions and documenting improvement, and communicating relevant information to the organizationwide quality assurance program (326).

[9] JCAHO recently modified its requirements to encourage the use of indicators from the clinical literature (326).

[10] Their physicians may be on the medical staff of a particular hospital and may direct the patient to that hospital so they may care for the patient there.

[11] The Accreditation Association for Ambulatory Health Care also has standards for emergency services, but their standards are oriented primarily toward freestanding emergency service centers (4). The ACEP/ENA guidelines apply to both hospital and freestanding emergency facilities (23).

regulatory bodies" (23). Similarly, AMA guidelines for emergency services, currently under revision, are to be considered guidelines for use by hospitals, rather than standards (36,38,178). Neither ACEP/ENA nor the AMA has any plans to survey for compliance with the guidelines they have devised. ACEP/ENA, American College of Surgeons, and AMA guidelines have, however, been adopted or adapted by some State bodies for regulatory use.

The following discussion makes a distinction between emergency services and trauma centers, but in practice, the distinction is not always clear. The medical services being evaluated in the trauma literature are not always restricted to trauma care (543), and there is some overlap in the guidelines for emergency services and trauma centers, as there is in the services themselves. Some trauma centers have their own admitting areas and staff (621), while others are a "concept" within the emergency department (153). In general, however, emergency medical services focus on prehospital care and care within the emergency department; trauma care includes inhospital and rehabilitative care.

Standards and Guidelines for Emergency Services.—Standards and guidelines for emergency services and standards can be distinguished along at least three dimensions: 1) whether they are standards or guidelines, 2) their breadth or depth, and 3) whether they distinguish among levels of services. Table 9-2 shows how various organizations' standards and guidelines for emergency services can be characterized along these dimensions.

The proposed AMA guidelines for emergency services will have perhaps the largest breadth, because they will be a compilation of guidelines from about 10 specialty organizations. The list of specialty organizations consulted in the development of the proposed AMA guidelines is shown in table 9-3.

Because the AMA guidelines for emergency services will incorporate the standards of specialty organizations, they will also have the greatest depth. The guidelines for emergency services of specialty organizations such as the ACEP/ENA, for example, designate administrative and managerial responsibilities, staffing levels, equipment. drugs, and relationships among the emergency service and other hospital departments (23). The ACEP/ENA guidelines do not specify guidelines for care of specific conditions such as burns or poisonings. Although the ACEP/ENA guidelines do not require that emergency departments operate continuously and do not stipulate levels of emergency care, they state that the emergency department should be staffed by a physician during all hours of operation. Optimally, according to the ACEP/ENA guidelines, the medical staff should be board certified in emergency medicine and the nursing staff should practice in accordance with the Standards of Emergency Nursing Practice.

Like the other organizations, JCAHO lists various aspects of hospital emergency services: organization, direction and staffing; integration, training and education, policies and procedures; and facility design and equipment. JCAHO standards for emergency services are more specific than

Table 9-2.—Characteristics of Various Organizations' Standards and Guidelines for Emergency Services

Organization	Standards or guidelines	Breadth v. depth	Levels of care
JCAHO	Standards[a]	Breadth	Levels I (highest) to IV (lowest)
HCFA	Standards[a]	Breadth	None specified
ACEP/ENA	Guidelines[b]	Breadth	None specified
AMA	Guidelines[b]	Breadth and depth	To be specified

Abbreviations: ACEP/ENA = American College of Emergency Physicians and Emergency Nurses Association; AMA = American Medical Association; HCFA = Health Care Financing Administration; JCAHO = Joint Commission for the Accreditation of Healthcare Organizations.

[a] These guidelines apply to hospitals only.

[b] These guidelines apply to freestanding emergency facilities as well as hospitals.

SOURCE: Office of Technology Assessment, 1988.

Table 9-3.—Specialty Organizations To Be Consulted in Developing the American Medical Association's "Guidelines for Classification of Hospital Emergency Capabilities," January 1988

Type of emergency	Organization providing guidelines or guidance
General medical	American College of Emergency Physicians[a]
Behavioral and psychiatric	American Psychiatric Association
Burn	American Burn Association
Cardiac	American College of Cardiology and American Hospital Association
Pediatric	American Academy of Pediatrics
Perinatal	American College of Obstetrics and Gynecology and American Academy of Pediatrics
Poisoning or drug	American Association of Poison Control
Spinal cord	American Spinal Cord Injury Association
Trauma	American College of Surgeons
Pediatric trauma	American Pediatric Surgery Association; American College of Surgeons

[a] Tentative.

SOURCE: P. Dietz, Program Administrator, Commission on Emergency Medical Services, American Medical Association, Chicago, IL, personal communication, Jan. 28, 1988.

Table 9-4.—HCFA's Condition of Participation Governing Emergency Services

482.55 *Emergency Services*
The hospital must meet the emergency needs of patients in accordance with acceptable standards of practice.

a. *Standard: Organization and direction.*
If emergency services are provided at the hospital:
1. The services must be organized under the direction of a qualified member of the medical staff; and
2. The services must be integrated with other departments of the hospital.
3. The policies and procedures governing medical care provided in the emergency service or department are established by and are a continuing responsibility of the medical staff.

b. *Standard: Personnel.*
1. The emergency services must be supervised by a qualified member of the medical staff.
2. There must be adequate medical and nursing personnel qualified in emergency care to meet the written emergency procedures and needs anticipated by the facility.

SOURCE: U.S. Department of Health and Human Services, Health Care Financing Administration, "Appendix A: Interpretive Guidelines and Survey Procedures—Hospitals," *State Operations Manual, Provider Certification*, HCFA-Pub. 7 (Baltimore, MD: September 1986).

ACEP/ENA guidelines with respect to the components of medical records for emergency patients and include requirements for quality control and monitoring and evaluation. In addition, JCAHO hospital-wide standards (e.g., medical staff requirements) apply to emergency services. Unlike the ACEP/ENA guidelines, JCAHO standards require that a hospital's emergency service be classified according to four levels of services provided, ranging from a "comprehensive" level of care (Level I) to a "first aid/referral" level of care (Level IV). The primary distinguishing feature among the four levels of emergency services is physician availability, although there also are differences with respect to nursing staff and equipment.

HCFA's condition of participation governing emergency services is rather broad (see table 9-4). They do, however, contain some of the same basic requirements as do the standards and guidelines of other groups. These requirements pertain to organization and direction and the qualifications of personnel. HCFA does not require specific staff coverage, equipment, or drugs.

Standards and guidelines for emergency services differ in their requirements regarding physician services. ACEP/ENA guidelines for emergency care recommend that emergency facilities be staffed during all hours of operation by a physician trained and experienced in emergency medicine. According to ACEP/ENA, unless there is physician staffing, a hospital should not be regarded as able to provide emergency services (709). This is a somewhat controversial recommendation. Not all of the 77 million visits to emergency facilities in a year (506) require a physician trained and certified in the specialty of emergency medicine, or even a physician. The basis of this ACEP/ENA guideline, however, is that "emergency health care exists for the individual benefit of the patient or family who perceives a need for emergency care, and for society's benefit in most casualty accidents" and that "the American public justifiably expects an emergency facility to be staffed by medical, nursing, and ancillary personnel who are trained and experienced in the treatment of emergencies" (23).

JCAHO's standards for emergency services do not require the presence of a physician at all times. JCAHO's standard for Level IV, the least com-

Photo credit: American College of Emergency Physicians

External standards and guidelines for emergency services differ on whether a physician must be available at all times in hospital emergency rooms.

prehensive level of care, for example, is that the "emergency service offers reasonable care in determining whether an emergency exists, renders lifesaving first aid, and makes appropriate referral to the nearest facilities that are capable of providing needed services." There must be some mechanism for providing physician coverage at all times in Level IV emergency facilities, but the mechanism is to be defined by the medical staff of the hospital. That the standard does not require immediate availability is reflected in JCAHO standards for Level III and higher emergency facilities. Level III facilities, for example, are required to have at least one physician available to the emergency care area within approximately 30 minutes.[12] The impact of having a trained and experienced physician available in an emergency department at all times has not been evaluated, so the relative validity of these standards cannot be judged.

In addition, it is noteworthy that only one set of standards or guidelines for emergency services—JCAHO's—requires that a hospital have a provision for providing emergency care 24 hours a day, 7 days a week (325).

Trauma Center Designations.—A review by the Centers for Disease Control of mortality data for 1984 shows that unintentional injuries were the leading cause of "years of potential life lost" before the age of 65 (440). A large proportion of efforts to decrease the number of deaths caused by injury have focused on injury prevention, but considerable attention has also been directed to the designation and implementation of emergency medical service systems (e.g., 454, 455; the Federal Emergency Medical Services Systems Act of 1973 [revised in 1975, repealed in 1981[13]]). In an organized emergency system, some hospitals are designated as regional trauma centers, to which severely multiply injured individuals are brought for treatment.

Intuitively, one expects that treatment and outcome in trauma centers will be better than elsewhere because of the immediate availability of rapid transportation, highly trained field personnel and emergency physicians, modern diagnostic tools, and experienced trauma surgeons (543). The only current national guidelines for trauma centers have been devised by the American College of Surgeons' Committee on Trauma (26). The American College of Surgeons' guidelines incorporate resources for both prehospital and hospital care. For hospitals, the guidelines specify the desired characteristics for three levels of trauma care. The two highest levels (Levels I and II) have similar requirements for patient care; the highest level (Level I) has additional requirements for education and research in trauma. Level III trauma center hospitals serve communities that do not have all the resources usually associated with Level I or II institutions; Level III facilities must have a "maximum commitment to trauma care commensurate with resources." Thus, for example, a Level III hospital might have a surgeon and other personnel on call rather than in-house. Nonetheless, a Level III facility would be called a trauma center by the American College of Surgeons.

According to a recent survey by the American College of Surgeons, approximately 177 hospitals have Level I trauma centers, 138 of which are designated as Level I by some external body; the remainder are self-designations by hospitals themselves. About 157 hospitals have Level II trauma centers, 124 of which are so designated by some external body (127). Table 9-5 indicates that only 19 States designate trauma centers using either the guidelines of the American College of Surgeons or a modified version of those guidelines.

---

[12] Level I and II hospitals are required to have at least one physician experienced in emergency care on duty in the emergency care area at all times. In addition, in Level I hospitals, there must be in-hospital physician coverage by members of the medical staff or by senior-level residents for at least medical, surgical, orthopedic, obstetric/gynecological, pediatric, and anesthesiology services.

[13] The Federal Government devolved much of its leadership responsibilities to States by folding the Emergency Medical Services Systems Act program into the Preventive Health and Health Services block grant.

The availability of various surgical, as opposed to medical, personnel is a major requirement for meeting the American College of Surgeons' guidelines, although there are numerous other requirements as well (26). For example, the American College of Surgeons recommends that a trauma team be organized and directed by a surgeon. The surgeon-directed trauma team is to evaluate the patient initially, and a surgeon is to be responsible for the patient's overall care. A physician with special competence in care of the critically injured is to be a designated "member" of the trauma team, and is to continuously staff the overall emergency department, but not be the head of the trauma team. Although the need for surgeons to deliver most trauma care is generally acknowledged, there is some controversy about who should design and manage the overall service (44).

RELIABILITY OF THE INDICATOR

Accreditation schemes for hospitals overall and for particular services are, it is clear, highly variable. To a consumer interested in neonatal intensive care, certification by the State of Ohio for a particular level of neonatal intensive care would convey much more information than the fact that a hospital with a neonatal intensive care unit had received JCAHO accreditation. Similarly, to a person interested in cancer care, a hospital's membership in the Association of Community Cancer Centers or designation as a Comprehensive Cancer Center by the National Cancer Institute would not convey the same type of approval as would approval by the Cancer Commission of the American College of Surgeons. Overall, HCFA's certification process is not as rigorous as JCAHO's accreditation process.

Some States have developed specific requirements for hospitals to offer specific services, but the types of services under these regulations and the specific requirements differ across States. In California, for example, emergency services are considered a supplemental service and appear as such on hospital licenses and published information for consumers (113,345); New York is about to change a similar regulation to make emergency services a basic requirement (472).

At the level of the individual State standard, there is considerable variation, because States develop their standards through statute and regulation, and statutes vary across States. The reliability of the surveyors and the survey process may vary as well. Hospitals surveyed by JCAHO, for example, have complained that judgments regarding their compliance with the same standard may vary considerably between survey periods. In part, the variation is due to periodic revision by JCAHO of its standards, a necessity.

VALIDITY OF THE INDICATOR

Accreditation for scope of hospital services is not a single entity, and individual standards themselves may vary in the extent to which they have been validated. Optimally, perhaps, standards and guidelines for scope of services would be based on medical practice with systematically demonstrated efficacy. The problem, however, is that much of medical practice is not based on evidence from scientific studies (628). Decisions about the "best" staff, equipment, and organization for a particular service or a particular problem are often the result of clinical judgment. Thus, most standards have been developed through expert consensus.

Table 9-5.—Characteristics of Trauma Center Designations by State[a]

	Categorization of facilities			Components of State laws and regulations				
	Categorize trauma centers by level	Categorize emergency services by AMA guidelines	Categorize specialized centers[b]	System designated to bypass nontrauma centers	Legal authority embedded in legislation or regulation	Use survey team from outside the State[c]	American College of Surgeons (ACS) guidelines	Modified 1981 ACS guidelines
Alabama		AL						
Alaska								
Arizona								
Arkansas								
California	CA	CA	CA	CA	CA	CA		CA
Colorado								
Connecticut	CT	CT	CT					
Delaware	DE	DE	DE	DE	DE	DE	DE	
District of Columbia	DC			DC	DC	DC		DC
Florida	FL			FL				
Georgia				GA	GA	GA		GA
Hawaii		HI						
Idaho				ID	ID			ID
Illinois			IL					
Indiana		IN						
Iowa				IA		IA		
Kansas								
Kentucky				KY				
Louisiana								
Maine								
Maryland	MD	MD	MD	MD	MD	MD		MD
Massachusetts				MA		MA	MA	
Michigan		MI						
Minnesota								
Mississippi			MS	MS		MS	MS	
Missouri				MO	MO	MO		MO
Montana								
Nebraska								
Nevada								
New Hampshire				NH		NH	NH	
New Jersey	NJ			NJ	NJ	NJ	NJ	
New Mexico	NM		NM	NM	NM	NM		NM
New York	NY							
North Carolina	NC	NC	NC	NC	NC	NC		NC
North Dakota								
Ohio								
Oklahoma								
Oregon	OR	OR	OR					
Pennsylvania	PA		PA	PA	PA	PA		PA
Rhode Island				RI	RI	RI	RI	
South Carolina	SC	SC	SC	SC	SC	SC		SC
South Dakota								
Tennessee								
Texas								
Utah	UT	UT	UT	UT	UT	UT		UT
Vermont				VT	VT			
Virginia				VA	VA	VA	VA	
Washington								
West Virginia	WV			WV	WV	WV	WV	
Wisconsin								
Wyoming								

[a] Excludes trauma center designations by counties and cities, of which there are few.

[b] Specialized centers include such facilities as spinal cord injury centers and burn centers.

[c] Some States, for example, use survey teams from or suggested by the American College of Surgeons.

SOURCE: H. Champion (Chief Director, Critical Surgical Care Services, Medstar, Washington Hospital Center), and H. Tetes (Attorney at Law, Bricker & Eckler), "State Laws and Regulations," Washington, DC, 1988.

Expert consensus may be an appropriate basis for establishing standards and guidelines for hospitals overall and for particular conditions, services, and departments. For some services, however, groups of experts disagree with one another, either on the need to establish standards or on the content of the standards themselves. The consumer is then left with the puzzling question of which group of standards or guidelines is more valid.

The validity of particular standards and guidelines could be demonstrated with studies of relationships between standards or guidelines and good process and outcomes, determined post hoc. Some standards and guidelines, such as those for neonatal intensive care and trauma centers, have been subjected to some such study, but most have not. JCAHO's hospital accreditation standards for hospitals have been subjected to very little study, and HCFA's hospital certification standards subjected to none. The studies that have been conducted have had methodological problems. For the most part, they have relied on retrospective analysis and outcomes as criteria and have not been conducted by independent observers.

One significant problem, applicable to all standards and guidelines, is that the standards or guidelines may change over time, sometimes significantly (388), a situation that makes the results of studies conducted at one point in time not applicable to subsequent standards. Frequent changes in standards may, of course, be necessary to reflect changes in technology and medical practice.

Validity of Overall Hospital Accreditation

There has been little attempt to validate overall JCAHO accreditation as an indicator of the quality of care. An important factor limiting studies seeking to validate JCAHO accreditation is that accreditation is refused or withdrawn for so few hospitals that the mere fact of accreditation may not be very sensitive to variations in quality. The few studies of the validity of JCAHO accreditation as an indicator of the quality of care have yielded inconclusive or noncomparable results.

Hyman obtained the results of JCAHO surveys for New York City hospitals (312). Unexpectedly, Hyman found that publicly supported hospitals had better JCAHO contingency scores[14] than voluntary not-for-profit hospitals on 9 of 11 functions. Friedman analyzed the relationships between numbers of JCAHO contingencies and HCFA's 1984 hospital mortality data (237). The result was a very low, statistically insignificant correlation, but this result is not surprising given the problems with HCFA's measure of hospital mortality (see ch. 4). One internal JCAHO study found a high level of agreement among JCAHO senior clinical and administrative staff as to the significance of several categories of standards for ensuring quality patient outcomes, but actual outcomes or process criteria were not used as validation standards (572).

Because JCAHO accreditation means that hospitals will be certified by HCFA, HCFA is required by law to validate JCAHO's results (Subsection 1864(c) of the Social Security Act). Every year, HCFA requests that State surveyors survey a small sample of JCAHO-accredited hospitals, stratified to be representative of hospitals nationally. HCFA also asks State surveyors to investigate patient complaints that seem to have substance. The State surveyors perform JCAHO validation surveys for HCFA using the Medicare conditions of participation. If a State surveyor finds that a hospital has significant deficiencies that could affect the health and safety of patients, the hospital is placed under State surveillance until the deficiencies are corrected. The hospital is no longer deemed to meet the Medicare conditions of participation, and the State monitors the correction of any deficiency.

HCFA conducted the last published JCAHO validation survey in fiscal year 1983, and transmitted it to Congress in 1986 (639). In general, JCAHO hospitals were found to be in compliance with HCFA's requirements. Any conclusion that JCAHO standards are valid because of their compliance with HCFA's requirements, however, depends on the validity of HCFA's survey process, and that process has not been validated. In addition, the discrepancy rates that HCFA found between HCFA's deficiencies and JCAHO's contingencies would mean that 276 hospitals in any single year, and as many as 750 hospitals overall[15] would be out of compliance on some condition of participation.

---

[14] At the time Hyman collected his data, JCAHO was using the terminology "recommendations" rather than "contingencies."

One future source of information for developing and validating JCAHO (and HCFA) standards is JCAHO's Agenda for Change project (see app. D). This project is attempting to develop more valid and condition-specific standards, including clinical process and outcome indicators. A potential JCAHO clinical indicator for obstetrics, for example, is birthweight-specific hospital mortality rates; hospitals designating themselves as high level neonatal intensive care units may have to meet a minimum birthweight-specific mortality rate. This project is being pilot-tested now with a small sample (324). In addition, JCAHO is progressing with plans to revamp its structural indicators so that they reflect the characteristics of effective health care organizations.

Validity of Standards and Guidelines for Specific Services

Many of the available studies of the validity of trauma center designations as indicators of the quality of care are methodologically flawed. Those that rely in whole or in part on autopsy studies, for example, are biased in that not all deaths result in autopsies. Some studies use different sources of information to determine causes of death. In one study of the San Diego County Regional Trauma System, for example, the causes of deaths in trauma centers were taken from a trauma registry, but the causes of death in comparison hospitals were taken from autopsies (564).

Perhaps more important, most studies of trauma center designations tend to be uncontrolled; that is, they merely compare patient outcomes before and after implementation of a trauma system. Studies that merely compare outcomes before and after implementation of a trauma system do not take into account factors other than medical care that may be responsible for reducing death rates from trauma (543). These factors may include simultaneous changes, such as reductions in speed limits and enhanced enforcement of drunk driving laws. In studies of standards and guidelines for neonatal intensive care, most of the research has been done only on Level III neonatal intensive care units (194), and the validation standards have been outcome measures, primarily mortality. Plans are underway to conduct studies of neonatal intensive care units using process criteria for validation.

Standards for emergency services have not been subject to the same amount of study that trauma center designations have, perhaps because the scope of services in emergency rooms is so broad. A knowledgeable observer concluded that there is no dependable knowledge about interhospital differences in emergency department performance or about the sources and correlates of such differences; there is also no dependable knowledge about the factors and conditions that facilitate or hinder emergency department effectiveness (245, 246).

FEASIBILITY OF USING THE INDICATOR

If validated, compliance with external standards for scope of hospital services is potentially an extremely valuable and easily accessible indicator of the quality of care for consumers. Currently, JCAHO and the American College of Surgeons both provide hospitals with a certificate to post. JCAHO's certificate addresses overall hospital accreditation, not individual services. Detailed reports on the results of JCAHO surveys of hospitals would be more informative; but these results are for the most part not easily obtained.

---

[15] HCFA's 1983 validation surveys found that up to 15 percent of hospitals surveyed were not in compliance with HCFA standards, although they had been in compliance with JCAHO's standards. If the 15 percent noncompliance rate is multiplied by the total number of JCAHO-surveyed hospitals (5,000), the number of hospitals not in compliance with HCFA standards would be 750.

JCAHO releases to the public, on request, information about whether a hospital is accredited, is involved in an appeal of its accreditation, is nonaccredited, or holds no accreditation status. JCAHO also releases a hospital's accreditation history. It does not, however, reveal a hospital's contingency score or copies of the survey reports. The JCAHO survey reports may be available on request from individual hospitals and from those States that require hospitals to submit the detailed survey reports as a requirement for licensure. Some States make the survey reports available; New York, Pennsylvania, and Arizona are among them. Other States, including California and Illinois, do not release copies of the JCAHO survey reports. States that recognize JCAHO accreditation for State hospital licensure purposes and require a copy of the accreditation report from the hospital are listed in table 9-6. JCAHO survey reports are long and technical, and consumers may face problems in interpreting the information they contain. One problem is that the survey reports focus on what is wrong with the surveyed hospitals. Without reviewing survey reports of several hospitals, consumers would not be aware of how a particular hospital compared with other hospitals.

Results of HCFA's hospital surveys exist in several forms. HCFA constructs individual hospital facility profiles that indicate the types of deficiencies a hospital has had for past survey years, and the services and personnel available at the hospital, among other information (649). In addition, HCFA constructs a table comparing State, regional, and national deficiency patterns for each Medicare condition of participation (650). A table constructed in January showed that 5 (27 percent) of the 18 HCFA-inspected hospitals in one State had deficiencies in the area of licensure of personnel; this rate compares to 19 percent for the U. S. Department of Health and Human Services Region III and 13 percent for the Nation (650). Some of the information from HCFA is not easy to use, however. The individual facility profiles report deficiencies by code numbers. These code numbers are not the same as the information on the report for the State, region and Nation, which does include written descriptions of the HCFA conditions of participation. It is, however, easy to glean from the individual facility profiles the services available at the hospital, which could be an important source of information for consumers. Both HCFA reports are intended as internal management tools for HCFA, but must be made available to the public on request (249). As for the individual survey reports, copies of a report (form 2567) that includes both the surveyors' recording of deficiencies and the hospital's plan of correction, and copies of the original survey reports from which the deficiency portion of form 2567 is drawn, are available from State survey offices, which are required to release them to the public (249).

Some States publish information about hospital accreditation and certification overall, licensure for particular services, and other information. California, for example, will send consumers who request it a summary report on hospitals. Hospitals in California and New York State must post in a conspicuous place their licenses, which note the services that the hospital is permitted to provide (345,414).

The feasibility of using scope of service designations to indicate quality of care is affected by the tendency of hospitals to self-designate themselves as specialists in particular areas. Even some State approval of trauma centers is based on hospital self-designation. Consumers would have to be careful that a designation is based at least on the stipulation of independent observers that the hospital adheres to a set of standards; otherwise such a designation may not be a valid indicator of quality. The American Hospital Association

Table 9-6.—States That Require Copies of JCAHO Accreditation Reports From Hospitals

Arizona
California
Connecticut
District of Columbia
Florida
Georgia
Idaho
Illinois
Iowa
Kansas
Louisiana
Maine
Massachusetts
Minnesota
Mississippi
Montana
Nebraska
New Hampshire
New York
North Carolina
Pennsylvania
South Carolina
Utah
Wyoming

SOURCE: Joint Commission on the Accreditation of Healthcare Organizations, "State Project Status Report," Chicago, IL, Sept. 21, 1987.

currently publishes a guide indicating the facilities and services available at hospitals that participate in the association's survey, but these designations are based largely on hospital self-reports. The American Trauma Society also publishes a list of trauma centers based on self-designation.

Consumers also face the problem of conflicting sets of standards for the same service. For cancer care, for example, there will soon be standards from two organizations (the American College of Surgeons and the Association of Community Cancer Centers). Although these standards build on each other to some extent, their relative validity remains to be established.

Even if available and reasonably validated, however, accreditation and standards for scope of services rely on the ability of patients to "match" their condition with the service as described by the accrediting body. When a patient requires more than one service, the problem becomes even more complex. Even accreditations that seem relatively condition-specific may not be useful to a particular patient. Hospitals whose cancer programs are approved, for example, may be more successful with some types of cancer than others.

CONCLUSIONS AND POLICY IMPLICATIONS

The external standards and guidelines that have been promulgated for hospital services overall and for scope of hospital services have not been rigorously validated as indicators of quality of care. Clearly, however, it seems worthwhile for consumers to seek out hospitals that have been judged by independent experts to have the appropriate resources to provide care, either overall or for specific conditions.

Some accreditation/certification information is readily available to consumers (see box 9-A). Information on a hospital's JCAHO accreditation history, for example, is available from JCAHO. HCFA will provide information on the certification status of any of the approximately 1,400 hospitals it inspects, and the American College of Surgeons will provide a list of the cancer programs it has approved. HCFA-inspected hospitals' actual survey reports are available from State agencies that conduct the surveys on behalf of HCFA. Some States require that hospitals post a notice stating which services they are allowed to perform and others provide consumers with reports supplying such information.

Other information is in existence but is more difficult for consumers to obtain or interpret. JCAHO survey reports, which form the basis of JCAHO accreditation decisions, are an example. Such reports can provide more detailed information to consumers than the mere fact of JCAHO accreditation. To see the reports, consumers may have to approach the hospitals themselves and ask for the reports, although some States will provide consumers copies of JCAHO survey reports for individual hospitals. Some consumers may have trouble interpreting and comparing detailed survey reports, and may prefer to see summary judgments that compare hospitals along a range of scores. Although JCAHO computes overall contingency scores for hospitals and also evaluates whether hospital emergency services meet requirements for four levels of care, this information is not readily available to the public.

There are considerably more guidelines available for the internal, optional use of hospitals than there are standards applied by independent groups of observers. Although hospitals may diligently conform to such guidelines, consumers should be wary of hospitals that say they adhere to the principles of one group or another, when there is no independent evaluation of such compliance.

Several steps could be taken to address the existing problems of external standards for overall hospital accreditation and scope of hospital services as quality-of-care indicators, to make existing information available, to improve existing standards, and to develop new standards. Table 9-7 shows the status of existing standards and guidelines in terms of their validity and feasibility of use as indicators of quality.

Box 9-A.—Selected Sources of Information About Scope of Hospital Services

Type of information	*Organization, address, or telephone number*
JCAHO hospital accreditation history	Joint Commission on Accreditation of Healthcare Organizations 1-800-621-8007 (nationwide except Illinois) 1-800-572-8089 (Illinois)
JCAHO hospital survey reports	Available from States of New York, Pennsylvania, and Arizona May be available from individual hospitals
HCFA hospital survey reports	Available from State agencies that conduct surveys on behalf of HCFA
List of hospital cancer programs approved by the American College of Surgeons	Cancer Department American College of Surgeons 55 East Erie Chicago, IL 60611

With some effort, existing information about compliance with existing standards could be made available to consumers. It seems ironic, for example, that survey reports for the 1,400 predominantly small and rural hospitals surveyed by HCFA are available to the public, while survey reports for the 5,000 hospitals surveyed by JCAHO on HCFA's behalf are not. Hospitals accredited by JCAHO are paid on the same basis as those certified more directly by HCFA. HCFA could improve its individual facility profiles, so that the reasons for deficiencies are intelligible to consumers and comparable to the reasons in HCFA's State, regional and national reports. As another example, JCAHO could include as part of its accreditation certificate the level of emergency services provided at a hospital, so that consumers could know whether a physician was likely to be on site. JCAHO and HCFA could develop summaries of their hospital survey reports that are meaningful to consumers (e.g., they could devise summary scores for specific services). Such information could be made available at hospitals themselves and in public places, such as libraries, local government offices, Social Security offices, and the offices of utilization and quality control peer review organizations. Similar information about the approvals by professional specialty organizations could also be made available.

Research to validate existing standards and help develop new standards is essential if consumers and providers are to be able to have confidence in the standards. Research is needed on all the standards and guidelines for scope of hospital services discussed in this chapter: JCAHO hospital accreditation standards; HCFA hospital certification standards; and various organizations' standards and guidelines for neonatal intensive care units, cancer care, emergency services and trauma units. Undoubtedly, research is needed on other condition-specific services. As some of the organizations that have developed standards begin to gather data about the process and outcome of care in organizations in compliance with the standards, the opportunities to conduct such research will increase.

Even as standards are being validated, the Federal Government, State governments, and private organizations could take more interest in developing and encouraging the use of consistent sets of standards for specific services and conditions. This step could increase consumers' access to scope of services information, as well as to hospitals with at least a minimal level of resources for conditions affecting them. Some consumers do not have access to scope of services information, because available guidelines are not applied

Table 9-7.—Characteristics of External Standards and Guidelines for Hospitals: Overall Accreditation/Certification and Specific Services

	Validated	Voluntary or mandatory	Survey by independent observers	Publicly available	Ease of access to information
Overall hospital accreditation/certification:					
JCAHO	Some studies; generally not	Voluntary	Yes	Accreditation history	Accreditation history easy; other information difficult
HCFA	No	Mandatory for participating hospitals	Yes	Yes	Difficult
Standards and guidelines for specific services:					
Neonatal intensive care services:					
AAP/ACOG	Level III/outcome studies	Voluntary	No	No	Difficult
States	Results differ by State	Varies	Ohio, some other States; not by AAP/ACOG	Ohio, some others	Difficult
Cancer care:					
ACS	No	Voluntary	Yes	Yes[a]	Fact of approval relatively easy; more detailed information difficult
ACCC	No	Voluntary	To begin	Yes[b]	Fact of membership relatively easy; actual adherence to standards difficult
Emergency services:					
JCAHO	No	Mandatory for participating hospitals	Yes	No, except through some States and willing hospitals	Difficult
ACEP/ENA	No	Voluntary	No	No	NA
AMA	No	Voluntary	No	No	NA
States	No	Both	?	Some	Varies
Trauma:					
ACS	No	Voluntary	Under consideration	No	NA
States	Some studies but poor methodologically	Both	Some	Probably	Varies

Abbreviations: AAP/ACOG = American Academy of Pediatrics and American College of Obstetricians and Gynecologists; ACCC = Association of Community Cancer Centers; ACEP/ENA = American College of Emergency Physicians and Emergency Nurses Association; ACS = American College of Surgeons; AMA = American Medical Association; HCFA = Health Care Financing Administration; JCAHO = Joint Commission on the Accreditation of Healthcare Organizations.

[a] List of approved hospitals.

[b] List of member hospitals, who may or may not follow ACCC guidelines.

SOURCE: Office of Technology Assessment, 1988.

by the organizations that developed them or by regulatory bodies. The development of standards has been slowed by professional rivalries, as well as by financial concerns (194,627), and lack of evidence about which requirements are valid. Less than half of all States have designated a State trauma center program, for example. Concerted efforts to develop consistent standards will be needed to overcome these problems.

In conclusion, considerable research is needed to validate accreditations/certifications for hospitals overall and external standards for specific hospital services. At present, accreditations, certifications, and approvals by independent bodies of experts seem to be a necessary, but not sufficient, indicator of a minimum standard of quality for hospitals overall and for some specific services. At the same time that research to develop more valid standards is being conducted, State and Federal governments could encourage the use and dissemination of information about hospitals' compliance with existing standards.

Chapter 10

Physician Specialization

CONTENTS

Page
Introduction 209
Reliability of the Indicator 212
Validity of the Indicator 214
  Is Physician Specialization a Reasonable Indicator of Quality? 214
  Does the Board Certification Process Accurately Reflect a Physician's Competence? 214
  Does Physician Specialization Accurately Predict a Physician's Quality? 215
  Is the Use of Physician Specialization as a Quality Indicator Generalizable Across Specialties? 221
Feasibility of Using the Indicator 222
Conclusions and Policy Implications 224

Tables

Table *Page*
10-1. Specialty Boards Recognized by the American Board of Medical Specialties 210
10-2. Specialty Boards Recognized by the Advisory Board for Osteopathic Specialists 211
10-3. Independent Boards That Certify Physicians 212
10-4. Recertification by Specialty Boards Recognized by the American Board of Medical Specialties: Current Status and Requirements 213
10-5. Studies on Physician Specialization Reviewed by OTA 216
10-6. Self-Designated Practice Specialties Recognized by the American Medical Association 223

Chapter 10

Physician Specialization

INTRODUCTION

The use of physician specialization to measure the quality of care provided by individual physicians represents a structural approach to measuring quality. Like other structural indicators, physician specialization is often used to assess quality on the assumption that certain characteristics of physicians may lead to better performance, which in turn may bring about better patient outcomes.

A person who wants to practice medicine and surgery legally in a State must obtain a license or certification of qualification from the State Board of Medical Examiners or other designated agency (70 Corpus Juris Sec. 12). Although the requirements for medical licensure vary among States, in general, a person must be a graduate of a medical school accredited by the Liaison Committee on Medical Education,[1] have completed 1 year of residency training in a program approved by the Accreditation Council for Graduate Medical Education,[2] and have passed the Federation Licensing Examination sponsored by the Federation of State Medical Boards (470).[3] With a medical license from a given State, a physician can practice medicine in that State, in whatever specialty area he or she chooses.

Some physicians, in addition to having general medical training, may have received training in a particular specialty area. Such training is not required for medical licensure, but physicians who have specialty training may be eligible to become certified by a specialty board.[4] Even if they have not received specialty training or been board-certified, however, physicians may designate themselves specialists.

Two major operational definitions of physician specialization have been used:

- certification by a specialty board, and
- the fact that a physician is practicing in his or her area of specialty training.

Many organizations certify physicians. The American Board of Medical Specialties and the American Medical Association (AMA) officially recognize the 23 specialty boards shown in table 10-1. These boards certify 63.5 percent of the physicians practicing in the United States (365). The Advisory Board for Osteopathic Specialists recognizes the 17 osteopathic specialty boards shown in table 10-2. All of the 40 specialty boards recognized either by the American Board of Medical Specialties or by the Advisory Board for Osteopathic Specialists require physicians to complete a specified amount of training and a certain set of requirements and to pass an examination. In addition to these boards, there exist at least 69 specialty boards not recognized by the American Board of Medical Specialties or the Advisory Board for Osteopathic Specialists (see table 10-3).

---

[1] The Liaison Committee on Medical Education is the official accrediting body for educational programs leading to the M.D. degree and is listed for this purpose by the U.S. Secretary of Education and recognized by the Council of Postsecondary Accreditation. The committee consists primarily of members from the Council on Medical Education of the American Medical Association and the Association of American Medical Colleges (157).

[2] The Accreditation Council for Graduate Medical Education is composed primarily of members from the American Board of Medical Specialties, the American Hospital Association, the American Medical Association, the Association of American Medical Colleges, and the Council of Medical Specialty Societies. Louisiana, Missouri, Ohio, and Tennessee do not require any residency training for licensure. Connecticut, Guam, Maine, New Hampshire, and Washington require completion of 2 years of residency training, and Nevada requires 3 years (470).

[3] Most States also recognize certifying examinations of the National Board of Medical Examiners to license physicians (513).

[4] Depending on the specialty, a physician may complete 1 to 5 years of additional training in a specialty area. The American Board of Orthopaedic Surgery requires 5 years of additional specialty training for a physician to become board certified, while the American Board of Colon and Rectal Surgery requires only 1 year of additional training. The term "board eligible" is sometimes used to describe a physician who has completed the necessary training and other predetermined requirements to become board certified, but has not taken the formal examination offered by the board. Because of continuing confusion about the term board eligible, however, the American Board of Medical Specialties' policy has disavowed the use of the term. The American Board of Medical Specialties has declared that the term has been given "such diverse meanings by different agencies that it has lost its usefulness as an indicator of a physician's progress toward certification by a specialty board" (18).

Table 10-1.—Specialty Boards Recognized by the American Board of Medical Specialties

American Board of:	General certification	Certificates in subspecialty areas		Date initial subspecialty offered
		Certificates of special qualifications	Certificates of added qualifications	
Allergy and Immunology	Allergy and immunology	Diagnostic laboratory immunology		1986
Anesthesiology	Anesthesiology	Critical care medicine		1986
Colon and Rectal Surgery	Colon and rectal surgery			
Dermatology	Dermatology	Dermatopathology		1974
		Dermatological immunology/ diagnostic and laboratory immunology		1985
Emergency Medicine	Emergency medicine			
Family Practice	Family practice		Geriatric medicine	1987
Internal Medicine	Internal medicine	Cardiovascular disease		1941
		Critical care medicine		1987
		Diagnostic laboratory immunology		1986
		Endocrinology and metabolism		1972
		Gastroenterology		1941
			Geriatric medicine	1988
		Hematology		1972
		Infectious disease		1972
		Medical oncology		1973
		Nephrology		1972
		Pulmonary disease		1941
		Rheumatology		1972
Neurological Surgery	Neurological surgery		Critical care medicine	
Nuclear Medicine	Nuclear medicine	Cooperates with American Board of Radiology and American Board of Pathology in radioisotopic pathology and nuclear radiology		
Obstetrics and Gynecology	Obstetrics and gynecology		Critical care	
		Gynecologic oncology		1974
		Maternal and fetal medicine		1974
		Reproductive endocrinology		1974
Ophthalmology	Ophthalmology			
Orthopaedic Surgery	Orthopaedic surgery		Hand surgery	
Otolaryngology	Otolaryngology			
				1973
Pathology	Anatomic and clin. path.	Blood banking		1950
	Anatomic pathology	Chemical pathology		1974
	Clinical pathology	Dermatopathology		1959
		Forensic pathology		1952
		Hematology		1983
		Immunopathology		1949
		Medical microbiology		1947
		Neuropathology		1974
		Radioisotopic pathology		1986
Pediatrics	Pediatrics	Diagnostic laboratory immunology		1961
		Pediatric cardiology		1987
		Pediatric critical care medicine		1978
		Pediatric endocrinology		1974
		Pediatric hematology-oncology		1974
		Pediatric nephrology		1986
		Pediatric pulmonology		1975
		Neonatal-perinatal medicine		
Physical Medicine and Rehabilitation	Physical medicine and rehabilitation			
Plastic Surgery	Plastic surgery		Hand surgery	
Preventive Medicine	Aerospace medicine Occupational medicine Public health and general preventive medicine			
Psychiatry and Neurology	Psychiatry	Child psychiatry		1959
	Neurology			
	Neurology with special qualifications in child neurology			
Radiology	Radiology	Nuclear radiology		1957
	Diagnostic radiology			
	Radiation oncology			
Surgery	Surgery		Hand surgery	
		Pediatric surgery		1975
			Surgical critical care	1986
		General vascular surgery		1982
			General vascular surgery	1988
Thoracic Surgery	Thoracic surgery			
Urology	Urology			

SOURCE: *American Board of Medical Specialties, Annual Report and Reference Handbook* (Evanston, IL: 1987).

Table 10-2.—Specialty Boards Recognized by the Advisory Board for Osteopathic Specialists

American Osteopathic Board of:	Subspecialties
Anesthesiology	
Dermatology	
Emergency Medicine	
General Practice	
Internal Medicine	Allergy/immunology Cardiology Endocrinology Gastroenterology Hematology Hematology/oncology Infectious diseases Medical diseases of the chest Nephrology Oncology Rheumatology
Neurology and Psychiatry	Child psychiatry Child neurology
Nuclear Medicine	
Obstetrics and Gynecology	Gynecologic oncology Maternal and fetal medicine Reproductive endocrinology
Ophthalmology and Otorhinolaryngology	Oro-facial plastic surgery Otorhinolaryngology and oro-facial plastic surgery
Orthopedic Surgery	Hand surgery
Pathology	Laboratory medicine Anatomic pathology Anatomic pathology and laboratory medicine Forensic pathology
Pediatrics	Neonatology Pediatric allergy/immunology Pediatric cardiology Pediatric hematology/oncology Pediatric infectious diseases Pediatric intensive care Pediatric nephrology
Preventive Medicine	Preventive medicine/aerospace medicine Preventive medicine/occupational-environmental medicine Preventive medicine/public health
Proctology	
Radiology	Diagnostic radiology Radiation oncology
Rehabilitation Medicine	
Surgery	Surgery (general) Neurological surgery Plastic and reconstructive surgery Thoracic cardiovascular surgery Urological surgery General vascular surgery

SOURCE: Advisory Board for Osteopathic Specialists, "Requirements for Certification: Advisory Board of Osteopathic Specialists and Boards of Certification," Chicago, IL, 1987.

In addition to offering a general certification, several specialty boards offer certificates in subspecialty areas. Altogether, there are 49 subspecialty areas of the 23 specialty boards recognized by the American Board of Medical Specialties (see table 10-1). Qualifications in these subspecialty areas are recognized by certificates of special or added qualifications.[5] Within the 17 specialty boards recognized by the Advisory Board for Osteopathic Specialists, there are 40 subspecialty areas (see table 10-2).

The 13 studies reviewed for this chapter pertain to board certification by the 23 specialty boards recognized by the American Board of Medical Specialties.[6] This chapter evaluates whether certification by these boards or practicing in one's area of specialty training are valid indicators of the quality of a physician's performance. Although the literature and this chapter examine physician specialization among allopathic physicians [Doctors of Medicine (M.D.s)], the discussion and conclusions drawn here are generally applicable to both allopathic and osteopathic physicians [Doctors of Osteopathy (D.O.s)].

The next two sections of this chapter evaluate the reliability and validity of physician specialization as a measure of the quality of care. The third section considers the feasibility of using physician specialization as a quality indicator. The final section of the chapter presents conclusions about physician specialization as an indicator of quality. That section also discusses methods to improve the reliability and validity of physician specialization as a quality indicator and considers alternatives to better assure consumers of the acceptability of a physician's quality of care.

---

[5] According to the American Board of Medical Specialties, "It is not necessary for physicians in a recognized specialty to hold special certification in a subspecialty of that field in order to be considered qualified to include aspects of that subspecialty within a specialty practice. Such special certification is a recognition of exceptional expertise and experience and has not been created to justify a differential fee schedule or to confer other professional advantages over other diplomates not so certified" (18).

[6] Additional details on the studies reviewed can be found in OTA's technical working paper "Physician Specialization as an Indicator of Quality: An Evaluation of the Literature" (434).

Table 10-3.—Independent Boards That Certify Physicians[a]

The following boards are called "American Board of ______" unless otherwise designated, and each claims to certify physicians.

Abdominal Surgeons
Acupuncture Medicine
Addictionology
Aesthetic Plastic Surgery
Alcoholism and Other Drug Dependencies
Algology (Chronic Pain)
Ambulatory Anesthesia
Bariatric Medicine
Bloodless Surgery
Chelation Therapy
Chemical Dependence
Clinical Chemistry
Clinical Ecology
Clinical Nutrition
Clinical Pharmacology
Clinical Toxicology
Cosmetic Plastic Surgery
Cosmetic Surgery
Council of Non-Board-Certified Physicians
Disability Evaluating Physicians
Electroencephalography
Electromyography and Electrodiagnosis
Epidemiology (College)
Facial Cosmetic Surgery
Facial Plastic Surgery
Forensic Psychiatry
Forensic Toxicology
Head, Facial & Neck Pain & TMJ Orthopedics
Health Physics
Homeotherapeutics
Insurance Medicine
Interventional Radiology
Laser Surgery
Law in Medicine
Legal Medicine
Malpractice Physicians
Maxillofacial Surgeons
Medical Accreditation (American Federation for)
Medical Genetics
Medical Hypnosis
Medical Laboratory Immunology
Medical Legal Analysis in Medicine & Surgery
Medical Legal & Workers Compensation Medicine & Surgery
Medical Legal Consultants
Medical Microbiology
Medical Preventics (Academy of)
Medical Psychotherapists
Medical Toxicology
Microbiology (Medical Microbiology)
Military Medicine
Neurological Orthopaedic Surgery
Nutrition
Otorhinolaryngology
Plastic Esthetic Surgeons
Prison Medicine
Psychiatric Medicine
Psychiatry (American National Board of)
Psychoanalysis (American Examining Board in)
Psychological Medicine (International)
Quality Assurance and Utilization Review
Radiology and Medical Imaging
Ringside Physicians and Surgeons
Skin Specialists
Spinal Cord Injury
Toxicology
Trauma Surgery
Tropical Medicine
Ultrasound Technology
Urologic Allied Health Professionals

[a] The boards listed below are not members of the American Board of Medical Specialties and are not recognized by the Advisory Board for Osteopathic Specialists.

SOURCE: American Board of Medical Specialties, "Self-Designated Boards," Evanston, IL, June 16, 1987.

RELIABILITY OF THE INDICATOR

Advances in medical science and technological changes are inherent in medical care. Unless a physician's knowledge and skills in a specialty area are periodically updated or assessed, physician specialization as represented by board certification or practicing in one's area of specialty training may be an unreliable measure of the quality of a physician's performance over time.

In the past 10 years, there has been an increasing trend towards recertification by specialty boards. The American Board of Medical Specialties encourages the periodic reassessment of physicians and has written guidelines on recertification for the specialty boards to use (18). So far, 15 of the 23 specialty boards recognized by the American Board of Medical Specialties have adopted or decided to adopt time-limited certification, and 1 board offers voluntary recertification (see table 10-4). Among these boards, the intervals between evaluations range from 6 to 10 years.

Without a recertification process, there is no guarantee that physicians have maintained the same level of skills and knowledge they demonstrated for their initial certification. Thus, board certification of a physician who was certified 20

Table 10-4.—Recertification by Specialty Boards Recognized by the American Board of Medical Specialties: Current Status and Requirements

American Board of:	Date of approval of recertification	Written exam for recertification	License required	Continuing medical education required	Recertification interval	Date of initial exam	Date of next recertification examination	Time-limited certificate[a]
Allergy and Immunology	1977	Yes		Yes[b]	6-10 yrs.	1977	1987	1989-10 yrs.
Dermatology							Indefinite	1991-10 yrs.
Emergency Medicine	1979				10 yrs.		1990	1980-10 yrs.
Family Practice	1969	Yes[c]	Yes	Yes	6 yrs.	1976	1987	1969-7 yrs.
Internal Medicine	1973	Yes	No	No	3 yrs.	1974	1987	1990[d]
Nuclear Medicine							Indefinite	1992-10 yrs.
Obstetrics/Gynecology	1976	Yes	Yes	No	10 yrs.	1978	1989	1986-10 yrs.
Ophthalmology	1986				10 yrs.		Indefinite	1992-10 yrs.
Orthopaedic Surgery	1980	Yes	Yes	Yes	6-10 yrs.	1983	1987	1986-10 yrs.
Otolaryngology	1986				7 yrs.	1995	Annually[e]	1989-7 yrs.
Pathology	1979	Yes[f]	Yes	Yes[g]	6-10 yrs.		Indefinite	No
Pediatrics	1977	Yes	Yes	No	6 yrs.	1980	1987	1988-7 yrs.
Phys. Med and Rehabilitation							Indefinite	1993-10 yrs
Plastic Surgery	1976	Yes	Yes	Yes	7-10 yrs.	1977	1992	1985-10 yrs.
Surgery	1978	Yes	Yes	Yes[b]	7-10 yrs.	1980	1987	1976-10 yrs.
Pediatric Surgery	1981	Yes	Yes	Yes[b]	7-10 yrs.	1982	1987	1075-10 yrs.
Thoracic Surgery	1979	Yes	Yes	Yes	7-10 yrs.	1984	Annually	1976-10 yrs.
Urology	1980	Yes	Yes	Yes	10 yrs.	1981	1987	1985-10 yrs.

[a] Beginning in the years indicated, boards have issued, or will issue, time-limited certificates.

[b] Type or number of hours unspecified.

[c] Also requires satisfactory review of ambulatory (office) records as a prerequisite to the recertification examination.

[d] Number of years not decided.

[e] After 1995.

[f] Required in one of three separate pathways.

[g] Requirement for two of three separate pathways.

SOURCE: American Board of Medical Specialties, *Annual Report and Reference Handbook* (Evanston, IL, 1987).

years ago may not indicate the same level of competence as the same board certification of a physician who has been assessed more recently.[7] This variability in the significance of board certification over time reduces the reliability of the use of board certification as an indicator of quality. Recertification requirements would increase the reliability of board certification as an indicator of quality by making its significance more constant over time.

A major issue in implementing recertification procedures is whether medical specialty boards can truly measure physician competence. Much of the opposition to recertification arises not because of recertification's quality assessment mechanism, but rather because of doubts about whether current examination procedures are an accurate measure of clinical skills. The American Board of Medical Specialties maintains that "recertification will be focused on performance assessment instead of the broad cognitive examinations used for primary certification" (365).

VALIDITY OF THE INDICATOR

Is Physician Specialization a Reasonable Indicator of Quality?

Intuitively, certification by a board recognized by the American Board of Medical Specialties is a valid indicator of the quality of a physician's medical performance. Board certification indicates that a physician has met a specified set of requirements and has performed up to a certain level on a qualifying examination in the specialty area. It makes sense that physicians who have had a certain amount of training in a specialty area would perform better than physicians who have had less or no training in the field.

Examples of the current uses of board certification demonstrate a general acceptance of its use as an indicator of quality. Patient brochures and other articles prescribing how to choose among physicians (published by hospitals or consumer health information centers) encourage consumers to use board certification as a measure of quality. Although the Joint Commission on Accreditation of Healthcare Organizations' standards for hospitals do not require that board certification be used for granting hospital privileges to physicians, the standards do state that "specialty board certification is an excellent benchmark for the delineation of clinical privileges" (330).

The fact that a physician is practicing medicine in the area in which that physician has been trained is also intuitively valid as an indicator of quality. It makes sense that a specialist practicing in the area in which he or she has been trained would provide better quality care than, say, a board-certified specialist who is not practicing in his or her area of specialization.

Does the Board Certification Process Accurately Reflect a Physician's Competence?

Many specialty boards limit their evaluations of a physician to evaluation of the physician's knowledge of the pertinent subject matter; they often do not evaluate a physician's interpersonal skills or skills used to technically apply their knowledge. Part of the reason for this situation may be that knowledge is fairly easy to test. By comparison, a physician's interpersonal skills are rather difficult to measure. Judgment and clinical skills are other qualities that are difficult to measure, yet are of utmost importance for determining the competence of a physician. The certification process of the American Board of Internal Medicine does include a form that asks several questions about the interpersonal skills a physician demonstrated during his or her period of residency (16). This form ("Evaluation Form for Clinical Competence") is sent to the program director of each physician's residency program.

At any rate, many of the aspects of a physician's practice mentioned in the last paragraph are only proxy measures of physician competence. Burg and Lloyd emphasize the importance of the specialty boards' defining competence to ensure the comprehensiveness of their evaluation measures:

---

[7] Physicians who were certified before their respective boards initiated recertification policies have not been subject to recertification.

> Definitions of competence within a medical specialty discipline serve the purpose of providing a first step toward the development of more valid procedures for the certification of specialists. This is because one form of validity, content validity, calls for a comprehensive delineation of the skills and abilities the board is attempting to measure. Ideally, measures of competence in a specialty should sample from the components of competence identified as important by members of that specialty (110).

To evaluate the competence of a physician with patients, it is important to use direct assessment methods. Evaluating a medical audit of pediatric performance, a study by the National Board of Medical Examiners demonstrated that for many common diagnoses, cognitive certifying exams do not test the same content area covered by direct audit assessments of clinical performance (612). Although several specialty boards utilize techniques in which a physician's practice is evaluated more directly by requiring information for specific cases treated, these methods have not been adopted by a majority of the boards.

Further complicating the issue of validity and board certification is that board-certified physicians' practices are not necessarily limited to the area in which they have been certified. In fact, statistics from the AMA's *Physician Characteristics and Distribution: 1986* demonstrate that 5 percent of board-certified physicians are not certified by the board corresponding to the primary area of their practice (35). This fact makes it impossible for board certification, which assesses a physician's knowledge and skills in one specialty area, to be an accurate reflection of all specialty areas in which a physician may practice. Physicians who are practicing in a specialty area in which they have not been board certified may not, outside of medical school or residency training, have had their skills assessed in that particular area.

Does Physician Specialization Accurately Predict a Physician's Quality?

A review of the 13 available studies on physician specialization and quality (see table 10-5) gives one little confidence that board certification accurately predicts which physicians will provide high-quality care and which will not (220,477, 481,604).

One explanation for board certification's low predictive power could be weaknesses in available studies. The studies may have too many methodological problems—small sample size, a biased physician sample (if it includes physicians' volunteering to participate), or no inclusion of patient-mix indices and severity-of-illness adjustments—to accurately assess the relationship between board certification and physician performance.

Other issues may affect the accuracy of board certification in predicting that physicians will provide high-quality care. If board certification were a mandatory process as opposed to a voluntary one, it would be more likely that non-board-certified physicians had failed the certification process and were substandard practitioners. Since board certification is voluntary, however, some physicians who are as well qualified as board-certified physicians may simply choose not to substantiate their training through certification. Of course, a percentage of non-board-certified physicians are physicians who have attempted and failed the certification process. Other non-board-certified physicians may not have met the board's mandated prerequisites to be eligible for certification. In the studies OTA reviewed, however, the percentage of unqualified non-board-certified physicians was not high enough to affect the performance results in favor of board-certified physicians.

Another possible explanation for board certification's not being predictive of a physician's performance in practice could be the imprecise evaluation procedures used by the specialty boards to certify physicians. All of the 23 boards recognized by the American Board of Medical Specialties use

Table 10-5.—Studies on Physician Specialization Reviewed by OTA

Study[a]	Physician specialties included	Conditions/ procedures studied	Level of aggregation	Performance measure	Adjustments for patient characteristics	Sample size	Results
Ramsey, et al., 1986 (510)	General internists	5 Conditions • Diabetes • Hypertension • Respiratory infection • Urinary tract infection • Ischemic heart disease	By diagnosis	Performance on American Board of Internal Medicine Exam Process measures • Items relating to diagnosis[b] • Treatment strategies[c] • Monitoring strategies[d] Outcome measures:[e] • Level of blood pressure control over a period of time • Glucose control • Exercise tolerance • Adverse outcomes[f] Patient satisfaction Evaluations by professional associates	No adjustments	185 Board-certified 75 Non-board-certified	• Board-certified physicians performed significantly better on certification exam. • No significant differences were found between board-certified and non-board-certified physicians for process or outcome measures for any condition. • No difference was found between mean patient satisfaction score for certified and non-board-certified physicians. Board-certified physicians received significantly higher ratings from professional associates in most categories.
Kelly and Hellinger, 1986 (347)	Primary surgeons: • Colon and rectal • Neurologic • Orthopedic • Thoracic • General	4 Conditions: • Stomach operation with cancer diagnosis • Stomach operation with ulcer diagnosis • Intestinal operation with cancer of the colon or rectum • Blood vessel surgery with abdominal aneurysm	By surgical procedure	Postsurgical mortality	Adjustments made for severity of illness, age, sex, and number of diagnoses	1,241 Total surgeons	Board-certified surgeons were found to be associated with lower patient mortality rates.
Strauss, et al., 1986 (604)	Pulmonary specialists General internists Family practitioners	Chronic obstructive pulmonary lung disease	By diagnosis	Outcome measures • Pulmonary function • Functional ability • Institutionalized days • Mortality	Adjustments made for severity and patient characteristics	96 Total physicians	No significant differences were found between the groups of specialists for outcome measures.
Goldberg and Dietrich, 1985 (257)	Family physicians General internists Medical subspecialists[g]	General primary care visits	By specialty status	Continuity of care[h]	Adjustments made for patient age, sex, and years with primary physician	40 Total physicians	No significant differences were found in continuity of care provided by subspecialists and generalists. Subspecialists provided higher levels of continuity to patients with a diagnosis lying within their areas of expertise, but only at high utilization levels.

Table 10-5.—Studies on Physician Specialization Reviewed by OTA (Continued)

Study[a]	Physician specialties included	Conditions/ procedures studied	Level of aggregation	Performance measure	Adjustments for patient characteristics	Sample size	Results
Sanazaro and Worth, 1985 (552)	General internists	5 Hospital diagnoses[i] 6 Office diagnoses[j]	By diagnosis	Physician evaluations by peer reviewers based on explicit criteria for substantiation of diagnosis, prescribing indicated drug regimen, and monitoring and attaining expected patient response	No adjustments	49 Board-certified 17 Non-board-certified	No significant differences were found between board-certified and non-board-certified physicians.
Payne, et al., 1984 (481)	Internists General/family practitioners Gynecologists Pediatricians	10 Conditions[k]	By diagnostic category	PPI/Physician Performance Score[l]	No adjusments	1,135 Total physicians	• Board-certification showed no effect on performance. • Specialists practicing in their area of training performed better than did family/general practitioners or specialists performing outside their specialty area.
Flood, et al., 1982 (220)	All specialties	15 Selected surgical procedures[m]	All surgical categories combined	Postsurgical status: measured extent of morbidity occurring 7 days after surgery or death within 40 days	Adjustments made for patient's physical status, stage of disease, age, sex, and emergency status at the time of the surgical procedure	544 Total physicians	Board-certification showed no effect on surgical outcomes.
Rhee, et al., 1981 (516)	18 Specialties[n]	19 Common diagnoses	Compared across diagnostic categories	PPI/Physician Performance Score[l]	No adjustments	321 Specialists 133 General practitioners	Specialists practicing outside their specialty areas had lower performance scores than specialists practicing in their domain. General practitioners practicing without limitation did not perform as well as specialists.
Lohr, et al., 1980 (385)	Internists Pediatricians Gynecologists General surgeons General practitioners	Common respiratory diagnoses and infectious diseases	By board-certification/specialty status	Appropriate use of injectable drugs and diagnostic tests (using explicit process criteria)	No adjustments	358 Total physicians (small sample size within each specialty)	Board-certified physicians had a higher rate of appropriate use of drugs and tests than non-board-certified physicians. No significant differences in appropriate use among primary care physicians or specialists. Poorer care was associated more with general practice and obstetrical gynecology; better care with pediatricians.
Garg, et al., 1979 (242)	Cardiologists Urologists Internists Family practitioners	5 Conditions: • Congestive heart disease • Ischemic heart attack • Stroke • Acute urinary infection • Chronic urinary infection	By diagnosis and specialty category	PPI	No adjustments	Not available	No differences were found between primary care physicians and subspecialists.

Table 10-5.—Studies on Physician Specialization Reviewed by OTA (Continued)

Study[a]	Physician specialties included	Conditions/ procedures studied	Level of aggregation	Performance measure	Adjustments for patient characteristics	Sample size	Results
Rhee, 1977 (515)	18 Specialties[n]	20 Diagnoses	Compared across diagnostic categories	PPI/Physician Performance Score[l]	No adjustments	321 Specialists 133 General practitioners	Board-certified/board-eligible physicians had higher performance scores than general practitioners or self-designated specialists.
Hulka, et al., 1976 (308)	Family/general practitioners Internists Pediatricians Obstetricians	4 Conditions: • Adult-onset diabetes mellitus • Congestive heart failure • Normal pregnant woman • Normal newborn during the first year of life	By diagnosis	Minimum explicit consensus criteria for management protocol	No adjustments	34 Family physicians 11 Internists 8 Pediatricians 8 Obstetricians	Pediatricians and obstetricians performed better for infancy and pregnancy. There were no differences among physicians in performance for diabetes mellitus or congestive heart failure.
Rhee, 1976 (514)	18 Specialties[n]	20 Diagnoses	Compared across diagnostic categories	PPI/Physician Performance Score[l]	No adjustments	321 Specialists 133 General practitioners	Physician specialists were not found to relate to quality of physician performance.

[a] Numbers in parentheses refer to numbered entries in the list of references at the end of this report.
[b] This includes determination of underlying factors, determination of severity of the condition, and determination of comorbid conditions.
[c] This includes complexity of therapy and avoidance of potentially harmful therapy.
[d] Frequency of followup office visits.
[e] Measured for chronic diseases.
[f] Hypotension, hypoglycemia, hypokalemia.
[g] Medical subspecialties included rheumatology, cardiology, hematology/oncology, and gastroenterology.
[h] Defined as the proportion of visits that patients received from their primary physicians.
[i] These diagnoses included chronic heart failure, acute myocardial infarction, chronic obstructive pulmonary diseases, diabetes mellitus, acute bacerial pneumonia.
[j] These diagnoses included hypertension, diabetes mellitus, chronic heart failure, angina pectoris, chronic obstructive pulmonary disease, osteoarthritis.
[k] Periodic adult medical examination, periodic gynecological examination, periodic pediatric medical examination, therapeutic use of chloramphenicol, keflex, digitalis preparation, and prednisone, anemia, essential hypertension, chronic heart disease (arteriosclerotic, hypertensive rheumatic), vulvovaginitis, acute urinary tract infection, chronic or recurrent urinary tract infection.
[l] The Physician Performance Index (PPI) is a process measure of performance developed by Payne and Lyons in 1984. Explicit process criteria for a variety of diagnoses and examinations were developed in 1974 by panels of practicing specialists. Physician performance was measured according to the level of physicians' compliance with these explicit criteria. The criteria were weighted by the physician panel so that a single PPI score for each diagnosis or examination was generated. A Physician Performance Score represents a physician's average PPI score over all of his or her treated cases.
[m] The surgical procedures included were gastric surgery for ulcer, selected surgery of the biliary tract, surgery of large bowel, appendectomy, splendectomy, abdominal hysterectomy, vaginal hysterectomy, craniotomy, amputation of lower limb (ankle to hip), repair of fractured hip, arthroplasty of the hip, lumbar laminectomy (with and without fusion), pulmonary resection, prostatectomy, and selected surgery of abdominal aorta and/or iliac arteries.
[n] Anesthesiology, dermatology, internal medicine, neurologic surgery, obstetrics/gynecology, ophthalmology, orthopedic surgery, otolaryngology, pathology, pediatrics, plastic surgery, preventive medicine, psychiatry and neurology, radiology, general surgery, thoracic surgery, urology, general practice.

SOURCE: Office of Technology Assessment, 1988.

written examinations with multiple-choice questions to evaluate physicians, and 16 of the boards require oral examinations. Nine of the boards require physicians to submit a case list for recertification.[8] In the oral examination for recertification, these nine boards ask physicians about their management of several cases. The particular cases discussed during the examination are picked by the specialty board from the case list submitted by the physician. Four boards require information from the physician's medical records (e.g., patient history, physician findings, and treatment outcome) for a specified number of cases (365).[9]

The relationship of the written and oral examination used by the American Board of Medical Specialties boards to actual physician performance is ambiguous. As noted earlier, test questions are more likely to measure what a physician knows about a certain field than the physician's actual clinical performance. If tests do not include an assessment of a physician's clinical competence, they may not provide an accurate prediction of a physician's performance. Studies have frequently demonstrated large discrepancies between levels of knowledge and levels of clinical performance (552).

Unfortunately, the assessment of a physician's clinical performance is not as adaptable to the format of a written examination as is an assessment of a physician's knowledge. Direct performance assessment instruments may provide a more accurate reflection of a physician's clinical competence and may have more predictive validity than written examinations. Although the 1975 American Board of Medical Specialties guidelines suggested that practice audits and performance evaluations should be part of the certification process, only four member boards—the American Boards of Family Practice, Obstetrics and Gynecology, Surgery, and Thoracic Surgery—have so far adopted such techniques. The American Board of Family Practice requires an office record review as part of its recertification process.[10] This board's particular methods of assessment increase the validity of its certification process, but they may give too much control to the physicians being evalu-

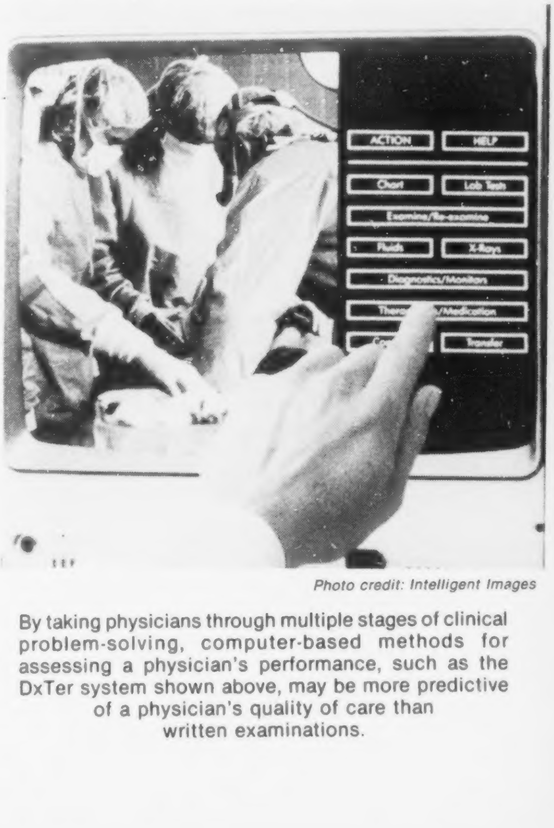


Photo credit: Intelligent Images

By taking physicians through multiple stages of clinical problem-solving, computer-based methods for assessing a physician's performance, such as the DxTer system shown above, may be more predictive of a physician's quality of care than written examinations.

---

[8] A case list specifies the number of diagnoses/procedures treated by a physician. The nine boards requiring case lists for evaluation are the American Boards of Colon and Rectal Surgery, Neurological Surgery, Obstetrics and Gynecology, Orthopaedic Surgery, Otolaryngology, Plastic Surgery, Surgery, Thoracic Surgery, and Urology.

[9] The four boards requiring information on cases treated are the American Boards of Obstetrics and Gynecology, Family Practice, Surgery, and Thoracic Surgery.

[10] Every 6 to 7 years, physicians certified by the American Board of Family Practice are required to undergo an office record review as part of their recertification process. This process involves each physician's choosing two individual patient records for each of three different conditions. The three conditions are chosen by the physician from a list of 20 possible conditions decided upon by the Board of Family Practice. The board sends the physician an extensive questionnaire and scansheet, to be filled out for each condition. Questions pertain to patient history, physical exam, medications prescribed, and diagnostic procedures. After receiving the completed scansheet, the board analyzes and scores it by computer. The scores are based on the physician's compliance with explicit process criteria (determined by the board) for the diagnosis and treatment of specific conditions. If the scansheets reveal that the physicians are not handling their patients as the board's standards dictate, the Board of Family Practice gives physicians the opportunity to send in additional patient records until their scansheets are approved. The board randomly selects physicians to be spot checked by requiring them to send in a specific patient record and comparing this record to the physician's own scansheet on that same patient. According to the Board of Family Practice, conflicts between a physician's scansheet and the spot checked records rarely occur (490).

ated. Physicians may be biased in their selection of records to send to the board for evaluation.[11]

Other approaches to office practice evaluation entail onsite reviews of actual patient records. The College of Family Physicians of Canada, which combines the functions of a certifying board and a professional association, has developed a medical practice quality assessment model based on chart abstractions (see ch. 7 for a description). The college plans to apply these techniques for use within certification examinations for "practice-eligible" candidates and for use in practice accreditation (85).[12] Similar onsite office evaluations are being performed by the College of Physicians and Surgeons of Ontario, the medical licensing body of Ontario,[13] and the Park Nicollet Medical Center in Minneapolis, Minnesota.[14]

Computer-based testing is another example of a technique to assess physician performance. Computer-based assessment techniques that reproduce a physician's clinical practice can provide an interactive representation of patient/physician encounters.[15] Although such techniques may be more predictive of a physician's quality of care than written examinations, their relative levels of accuracy remain to be validated through research.

The variations in the literature evaluating the relationship between board certification and quality of care reviewed for this OTA report are a reflection, in part, of the limited predictive power of various methods of assessing physician performance. The studies vary in regard to the performance measure used to assess a physician's quality of care. A physician's performance according to one technique may be different from his or her performance as measured by another method. In a study evaluating various procedures used to assess quality of care, Brook and Appel found the results of quality assessment to be determined by the method used (100). Between 1.4 and 63.2 percent of patients were determined to have received satisfactory care, depending on which method was used. Several studies reviewed for this report use the Physician Performance Index (PPI),[16] a process measure of performance developed by Payne and Lyons (242,481,514,515, 516). Other studies use different process measures or various outcome measures of performance (220,257,347,385,510,552,604).

Although there are several inconsistent results, the literature suggests that specialists practicing in the area in which they have been trained provide a higher quality of care than specialists practicing outside their area of training. Restricting their scope of practice presumably enables these specialists to treat patients in defined areas germane to their experience and training. One study evaluating performance differences between board-certified and non-board-certified physicians and between physicians practicing in their area of specialty training and physicians not practicing in their area of specialty training affirms the greater predictive power of physicians practicing in their area of training as an indicator of quality (481).

---

[11] If physicians choose only their best records, they will not be providing a representative sample of office records. Allowing a physician to send in records until the records are approved may also bias the sample.

[12] Practice accreditation, as opposed to certification, involves the random assessment of practicing physicians regardless of their specialty status.

[13] Each year since 1981, a total of 200 specialists and general practitioners out of all Ontario's physicians have been randomly selected to undergo a mandatory office evaluation, entitled the Peer Assessment Program. These evaluations are fairly subjective and basic in scope and structure. They are performed mainly for finding physicians with significant deficiencies in their records or patient care. Certain physician groups assumed to be more at risk of providing poor quality care, such as physicians over the age of 70 or those in solo practice, are specially targeted for review (139). Quebec's licensing body is involved in similar office reviews of physicians who are reported as needing special attention (54).

[14] A Primary Care Practice Profile was developed to supply information to the MedCenters Health Plan about the quality of care provided by family practitioners in various settings. The Primary Care Practice Profile is used by an audit team of three nurses and one physician during an onsite visit and incorporates a diagnostic-specific chart to evaluate medical records in physicians' offices (417).

[15] The National Board of Medical Examiners has developed a Computer-Based Exam that provides electronic patient simulation (456). The computer presents X-rays and electrocardiograms for example, and allows the physician to order any test or procedure required. The Computer-Based Exam keeps track of the results in terms of time, costs, and patient outcome. One physician score is produced. This exam has had the most experience, but other such patient simulation devices also exist. A DxTer system developed by David Allen of Intelligent Images allows the viewer to see the image of a person brought into an emergency room or clinical office (364). Although expensive to produce, this system provides a very realistic simulation of clinical practice. Another uncued system developed by Harless of the National Library of Medicine allows the physician to speak to the patient on the screen and receive a response from the patient (364).

[16] Explicit process criteria for a variety of diagnoses and examinations were developed in 1974 by panels of practicing specialists. Panels developing criteria for infectious disease, heart disease, hypertension, gynecology, and pediatrics were made up of specialists practicing in the corresponding fields. Physician performance was measured according to the level of physicians' compliance with these explicit criteria. The criteria were weighted by the physician panel so that a single PPI score for each diagnosis or examination was generated. Several studies are one step removed from comparing physicians' PPI scores, and go further to derive a performance score for each physician. A Physician Performance Score is calculated by taking a mean of the standard scores for all the cases a physician has treated.

Methodological weaknesses in available studies evaluating physicians practicing in their area of specialty training may explain some of the variations in the results of the studies (242,515,604). In some studies, the inclusion of self-designated specialists in the category of physicians practicing in their area of specialty training may have confounded the performance scores of these specialists, seemingly limiting the predictive power of practicing in one's area of training as an indicator of the quality of care.

The use of physician specialization as an indicator of quality is more valid when its meaning is clarified by either of the operational definitions used in this chapter than when its meaning is left undefined. As noted earlier in this chapter, physicians may designate themselves specialists regardless of the amount of training they have had in the field or whether they are currently practicing in that specialty. There are no regulations or guidelines that limit who may call themselves specialists. Because specialists are reimbursed by Medicare at a higher rate than nonspecialists, there are financial incentives for physicians to label themselves specialists even if they have not received specialty training. Consequently, physician specialization (unless a physician's training and qualifications can be identified with certainty) allows for a wide range of interpretations and is not an accurate predictor of quality.

Is the Use of Physician Specialization as a Quality Indicator Generalizable Across Specialties?

Board certification by one specialty board says little about a physician's qualifications in another specialty area. Owing to the variation in methods of practice across medical fields and to the wide range of certification techniques utilized by each of the 23 specialty boards that belong to the American Board of Medical Specialties,[17] board certification can be defined only as it applies to a specific board. Since many available studies assess only one type of board-certified specialist, their results address only that particular specialty. One study, for example, assessed only the performance of physicians certified by the American Board of Surgery as compared to surgeons not certified by that board (347).

Along the same lines, a major limitation with a number of the studies OTA reviewed is that they evaluate board-certified physicians or specialists practicing in their area of training in the aggregate rather than by specialty category. Aggregating the scores of physicians certified by different specialty boards and practicing in different specialty areas may mask subtle differences among specialty categories. Furthermore, specialists from each of the specialty groups may not be equally represented in the sample. One study with these problems compared board-certified and board-eligible physicians to general practitioners and self-designated specialists. Performance scores for the physicians were aggregated by board certification or eligibility status and by self-designation or general practice instead of being interpreted separately by specialty category (515).

There is an inevitable trade-off in generalizability between studies that assess physician performance with respect to many diagnoses and studies that assess performance with respect to only one diagnosis. Internal medicine, a specialty studied by Sanazaro and Worth (552), is an especially broad specialty field and includes many categories of subspecialties. An evaluation of subspecialists of internal medicine with respect to their performance of a subspecialty procedure, may be a more accurate approach to measuring the quality of a specialist's performance in a particular specialty, but would not necessarily be generalizable to other subspecialties.

The generalizability of physician specialization also has limitations within an area of practice. A physician's performance with respect to one diagnosis or procedure is not necessarily generalizable to other conditions and procedures. A certifying process that measures a physician's performance by assessing his or her skills for one or several diagnoses may not adequately serve as an indicator of the physician's performance overall (552). Thus, one cannot assume that a physician who performs poorly or well in the case of one or several diagnoses will perform poorly or well in all diagnoses.

---

[17] The boards may also differ in the required amount of time for education in a residency program and the specified stage for applying for certification—whether during or shortly after residency or after a specific amount of experience in the field.

In its broadest sense, certification by a board recognized by the American Board of Medical Specialties is an indication that a physician has met certain requirements and has passed certain examinations to practice in a specialty area. Unless board certification as an indicator of quality is used in reference to a specific specialty area, the generalizability of board certification as an indicator is unclear.

If one defines a specialist as a physician who is practicing in the area in which he or she has been trained, the link between previous training and specialty practice is set up. Unlike the board-certification literature, available studies on the quality of physicians practicing in their area of specialty training generally do classify specialization at the level of the individual specialty (242,257,385,481,604).

FEASIBILITY OF USING THE INDICATOR

Information on physicians' board-certification status and specialty designation is already available to individuals and organizations and easily understood by the public. *The American Medical Directory*, provided by the AMA, contains information on the American Board of Medical Specialties board certification, on self-designated primary and secondary practice specialties (see table 10-6), and on dates of recertification for all U.S. physicians alphabetically and geographically by city and county. This 4-volume directory is published every 2 years and is available in public libraries. *The ABMS Compendium of Certified Medical Specialists* is a 7-volume publication that lists all of the specialists certified by the 23 boards that belong to the American Board of Medical Specialties. Although this compendium is also published biennially, *The ABMS Compendium Supplement* is published in between publication dates of the original volumes and brings the lists of certified specialists more up to date.

The *Directory of Medical Specialists*, published by Marquis Who's Who and available in most public, hospital, medical, and university libraries, also contains information on board certification. This publication may be incomplete, however, because some boards do not supply information about new board-certified specialists to this source.

Data on American Board of Medical Specialties board-certification status and self-designated practice specialties are also available from the AMA Physician Masterfile. This computer data base contains current and historical information on all physicians practicing in the United States and on those U.S. physicians temporarily practicing overseas. The data are provided to the Masterfile by primary sources.[18] The AMA will provide a computerized printout containing background information on any U.S. physician to organizations such as hospitals, State licensing boards, medical schools, and medical societies for the purpose of credentials verification. Although information in the AMA Masterfile is primarily intended to assist organizations in verifying the credentials of physicians, the data are also available to individuals (672).

The general availability of information on board certification and specialty designation makes it apparent that confidentiality of this information is not an issue. Consumers may obtain information on board certification from county

---

[18] The American Board of Medical Specialties provides information on board certification, and the Federation of State Medical Boards (see ch. 6) provides information on final disciplinary actions by State boards that affect medical licensure.

Table 10-6.—Self-Designated Practice Specialties Recognized by the American Medical Association

Adolescent Medicine
Aerospace Medicine
Allergy
Allergy and Immunology
Anesthesiology
Cardiovascular Diseases
Critical Care Medicine
Dermatology
Dermatopathology
Diabetes
Diagnostic Laboratory Immunology
Emergency Medicine
Endocrinology
Facial Plastic Surgery, Otolaryngology
Family Practice
Gastroenterology
General Practice
General Preventive Medicine
Geriatrics
Gynecological Oncology
Gynecology
Hematology
Immunology
Immunopathology
Infectious Diseases
Internal Medicine
Legal Medicine
Maternal and Fetal Medicine
Medical Microbiology
Neonatal-Perinatal Medicine
Nephrology
Neurology
Neurology, Child
Neuropathology
Nuclear Medicine
Nuclear Radiology
Nutrition
Obstetrics
Obstetrics and Gynecology
Occupational Medicine
Oncology
Ophthalmology
Otolaryngology
Pathology, Anatomic/Clinical
Pathology, Anatomic
Pathology, Blood Banking
Pathology, Chemical
Pathology, Clinical
Pathology, Forensic
Pathology, Radioisotopic
Pediatrics
Pediatric Allergy
Pediatric Cardiology
Pediatric Endocrinology
Pediatric Hematology-Oncology
Pediatric Nephrology
Pediatric Pulmonology
Pharmacology, Clinical
Physical Medicine and Rehabilitation
Psychiatry
Psychiatry, Child
Psychoanalysis
Public Health
Pulmonary Diseases
Radiation Oncology
Radiology
Radiology, Diagnostic
Radiology, Pediatric
Reproductive Endocrinology
Rheumatology
Surgery, Abdominal
Surgery, Cardiovascular
Surgery, Colon and Rectal
Surgery, General
Surgery, Hand
Surgery, Head and Neck
Surgery, Neurological
Surgery, Orthopedic
Surgery, Pediatric
Surgery, Plastic
Surgery, Thoracic
Surgery, Traumatic
Surgery, Urological
Surgery, Vascular

SOURCE: American Medical Association, Division of Survey and Data Resources, "Intended Use of AMA Physician Masterfile Codes for Self-Designation of Practice Specialty," Chicago, IL, January 1987.

medical societies or from the American Board of Medical Specialties. Directly asking a physician or requesting the information from the hospital where the physician has staff privileges are further possibilities.

Determining whether physicians are practicing in the area of their training is not as easy for consumers as is determining a physician's board-certification status, largely because of inconsistencies regarding the qualifications and range of practice of a specialist. One study demonstrates a significant disparity between the number of physicians listed under the physician specialty headings in the Yellow Pages and the number of physicians listed in the *Directory of Medical Specialists* as board certified (511). In the current system, physicians can designate as their specialty an area in which they have had no or little training. Thus, consumers could be confused or misled by specialty designations. In most States except for Maryland, there is no system in place to substantiate that specialists have been trained in their practice areas.[19]

As noted earlier, requirements may vary substantially among the many boards that claim to certify physicians. For some boards, a set fee may be the only prerequisite for certification. A consumer uninformed about different types of certification may falsely assume that certification from a particular board is significant.

---

[19] In 1985, the Maryland General Assembly enacted legislation prohibiting physicians from presenting themselves to the public as specialists unless identified by the Maryland Board of Medical Examiners. Although any physician licensed in Maryland may apply for specialty designation, only physicians certified by the American Board of Medical Specialties are allowed automatically to publicly designate themselves as specialists. Other physicians must complete a form outlining specific training and experience in the requested specialty. The forms are reviewed for completeness and then referred to a multispecialty peer-review committee of the State medical society for evaluation. The Maryland Board of Medical Examiners makes the final decision on the basis of recommendations from the committee. The identification is permanent, and the physician may publicize himself or herself in that area of specialty (Annotated Code of Maryland, 10.32.09).

CONCLUSIONS AND POLICY IMPLICATIONS

Certification by the American Board of Medical Specialties or the Advisory Board for Osteopathic Specialists enables consumers to identify those physicians who meet a standard set of qualifications, including specific training in a specialty field and passing a certification examination. Consumers should be made aware, however, that such certification is not a reflection of the amount of practical experience a physician has in a specialty area or an adequate measure of demonstrated proficiency in the field.[20]

The low predictive power of current methods of assessing physician performance limits the validity of board certification as an indicator of high-quality care. Furthermore, board certification is not an indicator of quality that is generalizable across specialties, diagnoses, or procedures. An accurate measurement of a physician's performance requires that many diagnoses in the physician's specialty area be evaluated and interpreted individually.

Consumers should also be made aware that in most States, the fact that physicians designate themselves specialists does not necessarily mean that they have had advanced training in the specialty field corresponding to the area of their practice. To establish whether physicians are practicing in an area that correlates to their training, a consumer must verify that a physician's designated areas of practice match published listings of board-certification status. If a physician "specialist" is not board certified in any specialty field, there is no reliable method for a consumer to verify that the physician has had advanced specialty training.

To strengthen the validity of the relationship between board certification and clinical performance would require improving the reliability and validity of the specialty board evaluation procedures. The American Board of Medical Specialties recognizes this need and is involved in various studies working towards improving the predictive power of its specialty boards' evaluation processes (21).

The recertification of physician specialists would increase the reliability and the validity of certification. Recertification procedures could maintain the significance of board certification over time and would encourage physicians to update their skills. Although 16 boards recognized by the American Board of Medical Specialties have adopted some form of recertification, 7 boards still have not. Directories that list certification and recertification dates for specialists are publicly available, but consumers may not be aware of them (17,31).

Unfortunately, recertification would assure the public of the quality of care provided by only those physicians who hold a specialty certificate. Mandatory recredentialing (using valid performance assessment methods) for all licensed physicians would establish a more comprehensive system and would reassure the public of a physician's clinical competence.[21]

Performance assessments through medical audits would increase the predictive power of the certification/recertification process. The shift from a knowledge-based to a performance-based assessment, by making the process more a reflection of a physician's actual practice, would increase the validity of board certification. Langsley of the American Board of Medical Specialties writes, "Competence represents the potential for performance, but only performance assessment demonstrates that such potential is used in actual professional practice" (364).

---

[20] Although Medicare's conditions of participation for hospitals once used board certification as a requirement for participation, the conditions have since changed (FR 22021-22023, 1986). The current conditions of participation with respect to a hospital's medical staff require the hospital governing board to "ensure that under no circumstances is the accordance of staff membership or professional privileges in the hospital dependent solely upon certification, fellowship, or membership in a specialty body or society" (641).

[21] In a response to New York State Governor Mario Cuomo's proposal to require periodic reviews to measure physicians' clinical competence, an advisory committee agreed upon a plan to use examinations, peer review, or audits of office practice (similar to the systems used in Canada) to assess physicians who are not affiliated with a hospital (86). To determine the competence of physicians with hospital privileges, the committee is considering using hospital review systems that currently accredit physicians. A strong force behind this recredentialing initiative is a concern for the 10,000 to 15,000 non-hospital-affiliated physicians in New York who are subject to little peer review and critique (362).

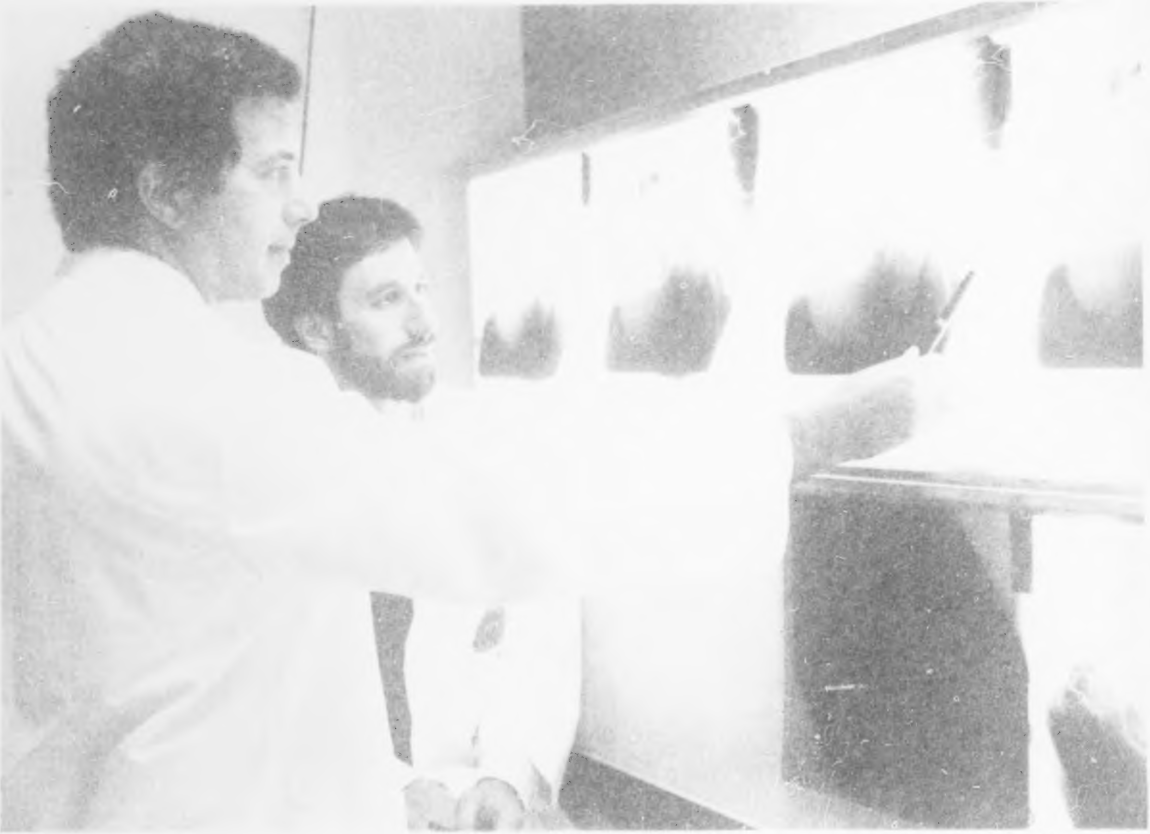


Photo credit: George Washington Medical Center

To improve the validity of physician specialization as an indicator of quality, State medical boards could be encouraged to prohibit physicians from representing themselves as specialists unless they demonstrate sufficient training and experience in the relevant specialty.

To prevent specialty designations from misleading consumers, there is need of a system to substantiate that physicians who represent themselves as being specialists have been trained in the specialty area in which they practice. A law enacted by the Maryland General Assembly prohibits physicians from representing themselves as specialists unless they demonstrate sufficient training and experience in the relevant specialty to the Maryland Board of Medical Examiners. Other State medical boards could be encouraged to implement similar systems to ensure that only physicians with special training are advertising themselves as specialists.

Physician specialization in conjunction with other criteria, such as years of practical experience or other verifications of a physician's clinical competence, would be a more significant indication of the quality of a physician's care than physician specialization alone. Most hospitals use a comprehensive credentialing process to delineate specific procedures that physicians are competent to perform depending on their qualifications. With a strong credentialing system in place, consumers can know that physicians on hospital staffs meet certain qualifications. Board certification or practicing in one's area of specialty training may be among the criteria in making this determination. Since these measures do not completely determine a physician's competence, however, hospitals also use other data to grant clinical privileges to physicians. The Joint Commission on the Accreditation of Healthcare Organizations requires that "professional criteria" be used for granting clinical privileges. These criteria include current licensure, relevant training or experience, current competence, and health status.

Perhaps a more rigorous system for evaluating specialized skills of physicians than the current process used for board certification would provide a more useful indication of the quality of a physician's care. The American College of Physicians, as a part of its Clinical Privileges Pilot Project, has developed guidelines defining minimum skills, education, and training that physicians need to perform competently eight specific medical procedures.[22] Although intended to assist hospitals by furnishing objective standards to assess clinical competence and delineate hospital privileges, these guidelines may be useful in assessing the competence of physicians performing these procedures on an ambulatory basis, where clinical review is less common.

Guidelines for granting physician privileges to perform specific procedures are also being developed by other organizations. The American Association of Urology set up guidelines for the training of physicians in the use of the extracorporeal shock-wave lithotripsy.[23] In addition, in 1987, a task force from the American College of Cardiology proposed training standards for performing coronary angioplasty.[24]

To date, delineated qualifications for physicians performing specific procedures are mostly in the form of guidelines. In actuality, any licensed physician could perform any number of complicated surgeries or medical procedures unrestrictedly. Many procedures are done at ambulatory clinics or "surgicenters," where there may not be a formal physician credentialing system in place. In these settings, there is a need to assure the public that the physicians performing these procedures are competent and well qualified. Perhaps the certificates of special competence currently offered by specialty boards could be used for this purpose. The qualifications needed by a physician to acquire such a certificate could be made more rigorous—instead of representing only the passing of an examination, perhaps representing a delineated amount of training, experience, and competence that the physician has acquired in performing the procedure. The qualifications required for a physician to obtain a certificate of special competence could be those credentials demonstrated to relate to better outcomes of patient care.

To encourage physicians to acquire the special training and experience to perform procedures, those physicians who hold such certificates could be reimbursed by Medicare at a higher rate. A more stringent regulation might be to require under Medicare's conditions of participation that a physician hold a certificate of special competence.

Additional research is needed to explore the qualifications of physicians that relate to improved quality of care. Research on specific procedures, similar to the studies of the American College of Physicians on eight specific medical procedures (24), is needed to determine an adequate standard of training and experience for physicians to perform the procedures.

Also needed is research that evaluates board certification and physicians practicing in their area of training by type of specialty. The conditions or procedures chosen to be evaluated within each study should be conditions that are highly prevalent among the types of patients a particular specialist sees, so that they can serve as a valid representation of a specialist's practice. Caution would still be warranted with respect to generalizing from a physician's performance for one condition to performance for other diagnoses.

---

[22] So far, guidelines have been developed for renal biopsy, acute hemodialysis, acute peritoneal dialysis, continuous arteriovenous hemofiltration, flexible fiberoptic sigmoidoscopy, colonoscopy, esophagogastroduodenoscopy, and endoscopic retrograde cholangiopancreatography. These guidelines, which have become official policy of the American College of Physicians, were based on medical literature and expert consensus. Approximately 4,200 general internists and subspecialists were surveyed as part of the pilot project to obtain more objective data about the experience and training necessary for competence in the procedures. A project to develop guidelines for a number of other procedures in all subspecialties of internal medicine is planned for 1988 (24).

[23] To obtain a certificate of training for this procedure, a physician is required to have had 5 days of training and to have performed 15 procedures. Most hospitals require that a physician be certified in the use of the lithotripter before he or she is allowed to perform the procedure. The Association also approves potential sites where lithotripsy training can take place (380).

[24] The task force calls for "three levels of training for three types of cardiologists: those doing cardiac catheterization or angiography, those doing both, and those doing both plus angioplasty or other advanced procedures that may be developed" (22). The guidelines state that the physician training to perform angioplasty must complete a fourth year of residency training and a minimum of 125 coronary angioplasty procedures, including 75 as primary operator. The task force also calls for a "certifying process for a certificate of added experience and qualification in advanced cardiac catheterization procedures [such as angioplasty]."

To increase the validity of various measures of physician performance, further research on various performance assessment techniques needs to be conducted. Techniques with greater predictive power would increase the significance of board certification as an indicator of the quality of a physician's performance. Validated methods to assess physician performance would also increase the significance of other criteria used in determining a physician's competence, such as the certificates of special competence. Consumers could then rely more heavily on these criteria as acceptable indicators of quality care.

Chapter 11

Patients' Assessments of Their Care

CONTENTS

Page
Introduction 231
Reliability of the Indicator 232
Validity of the Indicator 235
  Bias in Patients' Ratings 236
  Validity of Patients' Assessments of Ambulatory Care 237
  Validity of Patients' Assessments of Inpatient Care 241
Feasibility of Using the Indicator 243
Conclusions and Policy Implications 245

Tables

Table *Page*
11-1. The Distinction Between Patients' Ratings and Patients' Reports Regarding the Quality of Medical Care 232
11-2. Reliability of Patients' Ratings of Ambulatory and Inpatient Medical Care: Findings From Studies Reviewed by OTA 233
11-3. Number of Studies Reviewed by OTA on the Validity of Patients' Assessments of the Quality of Medical Care 236
11-4. Validity of Patients' Assessments of the Interpersonal Aspects of Ambulatory Care: Findings From Studies Reviewed by OTA 238
11-5. Validity of Patients' Assessments of the Technical Apects of Ambulatory Care: Findings From Studies Reviewed by OTA 240
11-6. Validity of Patients' Assessments of the Interpersonal Aspects of Inpatient Care: Findings From Studies Reviewed by OTA 241
11-7. Validity of Patients' Assessments of the Technical Aspects of Inpatient Care: Findings From the Study Reviewed by OTA 242
11-8. Validity of Patients' Assessments of the Overall Quality of Inpatient Care: Findings From Studies Reviewed by OTA 243

Chapter 11

Patients' Assessments of Their Care[1]

INTRODUCTION

Along with the current resurgence of interest in assessing the quality of medical care has come renewed attention to the patient's viewpoint and increasing efforts to involve patients in quality assessment activities. Although several factors may motivate decisions to involve patients in such activities, one assumption appears to be critical: that patients can provide valid information about the quality of their medical care.

Seeking input from patients when evaluating quality of medical care has at least three rationales. First, it ensures that evaluations will represent the values of the individual consumers of medical services. Second, patients are the only source of information regarding certain aspects of medical care (particularly the interpersonal aspects of the provider-patient relationship) and also may provide information that supplements information from other traditional sources, such as medical records (133,247,368,400,615,689). Thus, patients can provide both unique and supplementary data on attributes related to the quality of medical care. Third, patient surveys probably cost no more and may cost less than data obtained for quality assessment from other sources (161).

Conclusions about the validity of patients' assessments of the quality of medical care are likely to vary depending on the aspect of quality being evaluated. Patients are clearly more qualified to judge the interpersonal aspects of quality than the technical aspects.[2] Alternative sources of data may provide better information about the quality of the technical process of care. Furthermore, the amount of evidence regarding the validity of patients' assessments varies greatly depending on what aspect of quality is under consideration. The majority of studies have been done in ambulatory settings and have tended to focus on interpersonal aspects of care.

The evidence relevant to the validity of patients' assessments of ambulatory and inpatient care is discussed separately in this chapter. Because different aspects of the quality of care have been studied in inpatient and ambulatory settings, the measures used to test and define validity have also varied. Moreover, the state-of-the-art of defining concepts and developing and validating measures of quality is much further advanced for patients' assessments of ambulatory care than for those of inpatient care. Finally, unlike the literature on ambulatory settings, the literature on inpatient care lacks a coherent taxonomy of quality from the patient's perspective. Many patient-based indicators of quality used in inpatient settings represent considerable aggregation of various aspects of quality (i.e., assessment of quality in general). For that reason, this chapter's review of the evidence from inpatient settings includes discussions not only of interpersonal and technical aspects of quality but of overall quality as well.

Patients' ratings of their care must be distinguished from patients' reports about their care. *Patients' ratings* represent personal evaluations of aspects of medical care providers and services; because ratings reflect personal experiences, expectations, and preferences, as well as the standards patients apply when evaluating care, ratings are inherently more subjective than reports.[3] *Patients'*

---

[1] This chapter is based on a paper prepared for OTA by John E. Ware, Jr., Allyson Ross Davies, and Haya R. Rubin (686).

[2] The definition of the quality of medical care used in this report excludes most aspects of the availability and accessibility of care (see ch. 3). Nevertheless, evidence from five studies reviewed by OTA supports the validity of patients' assessments of access to care in the ambulatory setting. Patients' ratings of specific features of access, including resource availability, office and appointment waiting times, and waiting time for emergency treatment (8,13,162,378), as well as overall access (398), were significantly related to independently and objectively observed differences in these features of ambulatory care.

[3] Most of the studies reviewed here used attitudinal measures, or more specifically, patient satisfaction measures, to obtain data from consumers. The majority of items in these surveys can be considered evaluations, either because the respondents are asked the strength of their endorsements of an evaluative statement (e.g., "My doctors are very competent and well-trained"), or because the response categories offered constitute an evaluation (e.g., "excellent" to "poor"; "satisfied" to "dissatisfied").

reports deal with things that did or did not occur; they are inherently more objective than patients' ratings and can be more readily confirmed by an outside observer. Table 11-1 illustrates the distinction between patients' ratings of the technical and interpersonal aspects of care and patients' reports about these aspects.

For the literature review that was the basis for the analysis in this chapter, over 450 publications on the subject of patients' assessments of their medical care were screened, and 50 articles were selected for indepth review. In choosing studies, greatest emphasis was placed on studies that, from a methodological perspective, had strong designs and provided adequate information about interventions (if any), reliability of patient data, and operational definitions of the variables studied in relation to patient information. In particular, studies were favored that formally tested for a direct link between actual differences in the quality of care and patients' ratings or reports, either by manipulating quality experimentally or by obtaining measures of actual quality independent of those provided by patients. In the case of experiments, some manipulation of the physician-patient encounter was required to determine that quality had actually been altered.

Table 11-1.—The Distinction Between Patients' Ratings and Patients' Reports Regarding the Quality of Medical Care

Aspects of quality being evaluated	Patient rating	Patient report
Technical aspects	Evaluation (e.g., excellent, good, fair, poor) of completeness of physical exam	Indication (yes-no) of whether physician did throat swab
Interpersonal aspects	Evaluation (e.g., excellent, good, fair, poor) of physician's friendliness	Indication (yes-no) of whether physician referred to patient by name

SOURCE: Office of Technology Assessment, 1988.

This chapter analyzes the reliability, validity, and feasibility of using patients' assessments as an indicator of the technical and interpersonal aspects of ambulatory and inpatient care. Where empirical evidence of validity is sparse or lacking, the types of information that are needed are identified. The practical considerations involved in obtaining data from consumers for purposes of evaluating the quality of physician and hospital care are also addressed.

RELIABILITY OF THE INDICATOR

Estimates of reliability[4] can be obtained in various ways: 1) by correlating scores on two forms of a measure (alternate-forms), 2) by correlating scores for the same measure at two points in time (test-retest), or 3) by correlating scores on items measuring the same concept (internal-consistency). Whatever method is used, a reliability estimate ranges from 0.0 to 1.0. Generally, the minimum acceptable standard for reliability of measures used in group comparisons (e.g., patient samples from two outpatient clinics or two hospitals) ranges from 0.50 (294) to 0.70 (467). Most uses of patient information on quality-related topics will involve group, rather than individual, comparisons.

Despite the importance of reliability estimates, the majority of studies identified in OTA's literature review of patients' assessments did not report estimates. Nineteen of the studies of patients' ratings in ambulatory settings included reliability estimates, as did 8 studies of patients' ratings in inpatient settings; several studies reported estimates for more than one sample. In studies that did report reliability estimates, the estimates exceeded 0.50 for virtually all multi-item rating scales, and many exceeded the 0.70 standard (see table 11-2). This finding holds for many relatively short (fewer than 10 items per concept) but well-constructed multi-item measures, even in disadvantaged populations where reliability tends to be poorer (310,691). Although many single-item ratings do not meet this minimum reliability standard (687), recent work suggests that reliable single-item ratings can be constructed (688).

---

[4] Reliability is the proportion of measured variance that is the true score, as opposed to random error.

Table 11-2.—Reliability of Patients' Ratings of Ambulatory and Inpatient Medical Care: Findings From Studies Reviewed by OTA

Study[a]	Sample size	Method of estimating reliability	Dimension(s) of care being rated	Number of items used to measure dimension[b]	Reliability estimate
PATIENTS' RATINGS OF AMBULATORY MEDICAL CARE					
Franklin and McLemore, 1967, 1970 (228,229)	136	IC	General satisfaction	20 items	.87
Hulka, et al., 1970 (310)	49	AF	Total:	42 items	.80
			Professional competence	12-14 items	.63
			Personal qualities	12-14 items	.75
			Access/finances	12-14 items	.43
Zyzanski, et al., 1974 (722)..	426	IC	Total:	42 items	.90
			Professional competence	14 items	.75
			Personal qualities	14 items	.86
			Access/finances	14 items	.68
Rojek, et al., 1975 (528)	1,100	IC	General satisfaction	3 items	.71
Aday and Anderson, 1975 (7)	2,000	IC	Total:	11 items	.84
			Access/finances	3 items	.68
			Interpersonal/technical quality	8 items	.90
Ware, et al., 1975 (692).....	903	IC	Availability total	10 items	.83
			Continuity	4 items	.78
			Finances	4 items	.69
			Interpersonal/technical total:	25 items	.89
			Interpersonal	3 items	.67
			Technical quality	4 items	.89
Ware and Snyder, 1975 (690)	433	IC	Availability (4)[c]	2 items	.47-.76
			Accessibility (3)	2-3 items	.49-.64
			Continuity (2)	2 items	.57-.67
			Finances (3)	2 items	.66-.75
			Interpersonal (3)	3-4 items	67-.75
			Technical quality (5)	2-4 items	.52-.73
Ware and Snyder, 1975 (690)	167	TRT (6-week interval)	Availability (4)[c]	2 items	.57-.62
			Accessibility (3)	2-3 items	.59-.62
			Continuity (2)	2 items	.59-.64
			Finances (3)	2 items	.62-.69
			Interpersonal (3)	3-4 items	.62-.69
			Technical quality (5)	2-4 items	.64-.70
Roter, 1977 (540)	250	IC	Overall satisfaction	6 items	.67
DiMatteo, et al., 1980 (181) .	4 to 10 pts. per doctor, 29 doctors	IC (across pts. for individual doctor)	Interpersonal aspects (10 pts/doctor)	NA	.61
			Interpersonal aspects (4-5 pts/doctor)	NA	.12
Breslau and Mortimer, 1981 (92)	370	IC	Total:		.51-.82[d]
			Access	8 items	
			Availability of resources	5 items	
			Continuity	3 items	
			Finances	6 items	
			Humaneness	8 items	
			Technical quality	7 items	
Ware, et al., 1981 (684).....	2,287	IC	Technical quality	4 items	.70
			Interpersonal aspects	3 items	.66
			General satisfaction	4 items	.74
			Access to care	7 items	.74

Table 11-2.—Reliability of Patients' Ratings of Ambulatory and Inpatient Medical Care: Findings From Studies Reviewed by OTA (Continued)

Study[a]	Sample size	Method of estimating reliability	Dimension(s) of care being rated	Number of items used to measure dimension[b]	Reliability estimate
Marquis, et al., 1983 (406)	279	IC	General satisfaction	4 items	.70
Bartlett, et al., 1984 (56)	60	IC	Overall satisfaction	8 items	.88
Chang, et al., 1984 (128)	268	IC	Global satisfaction	7 items	.95
Corah, et al., 1984 (150)	24	IC	Total:	10 items	.84
			Information/communication	3 items	.94
			Understanding/acceptance	3 items	.87
			Technical competence	4 items	.84
DiMatteo, et al., 1986 (180)	329	IC	Communication	8 items	.72
			Affective care	9 items	.79
			Technical care	3 items	.65
DiMatteo, et al., 1986 (180)	6 to 7 pts. per doctor, 57 doctors	IC (across pts. for individual doctor)	Communication	8 items	.52
			Affective care	9 items	.46
			Technical care	3 items	.40
Cope, et al., 1986 (149)	424	IC	Art of care	9 items	.92
			Technical quality	5 items	.81
Davies, et al., forthcoming (163)	1,537	IC	Access total	8 items	.65-.70[e]
			Availability total:	5 items	.64-.74
			Avail. of family doctors	2 items	.73-.78
			Avail. of hospitals	3 items	.78-.84
			Costs of care	2 items	.71-.80
			Quality total:	16 items	.89-.91
			Interpersonal aspects	8 items	.80-.83
			Technical quality	6 items	.65-.72
			Facilities	2 items	.78-.83
			General satisfaction	4 items	.66-.75
Ware, et al., forthcoming (689)	109	IC	Interpersonal aspects	5 items	.93
			Technical quality	5 items	.90
PATIENTS' RATINGS OF INPATIENT MEDICAL CARE					
Rice, et al., 1963 (517)	457	TRT (1 week)	Ward Evaluation Scale:	69 items	.81
			Physical facilities	22 items	.78
			Patient service	27 items	.77
			Patient management	20 items	.67
Souelem, 1955 (585)	95	AF	General attitudes toward mental hospitals	36 items	.88
Hinshaw and Atwood, 1982 (298)	5 studies, n's ranged from 49 to 237	IC	Total:	NA	.89
			Technical/professional	NA	.64-.97
			Trusting relationship	NA	.82-.92
			Education	NA	.49-.95
Wales, et al., 1983 (679)	115	IC	Total:	NA	.88
			General	NA	.82
			Competency	NA	.73
			Humaneness	NA	.72
			Physical environment	NA	.80
Wales, et al., 1983 (679)	115	TRT (24-hour)	Total:	NA	.93,.91[f]
			General	NA	.87,.92
			Competency	NA	.85,.84

Table 11-2.—Reliability of Patients' Ratings of Ambulatory and Inpatient Medical Care: Findings From Studies Reviewed by OTA (Continued)

Study[a]	Sample size	Method of estimating reliability	Dimension(s) of care being rated	Number of items used to measure dimension[b]	Reliability estimate
			Humaneness	NA	.83,.83
			Physical environment	NA	.92,.86
Carmel, 1985 (120)	476	IC	Physicians	11 items	.94
			Nurses	11 items	.95
			Supportive services	9 items	.86
Greenley, et al., 1985 (265)	177	IC	Humaneness of staff	NA	.85
			Humaneness of psychiatrist	NA	.95
Casarreal, et al., 1986 (124)	972	IC	Admitting attitudes	3 items	.79
			Nursing attitudes	5 items	.82
			Physician attitudes	5 items	.85
			Housekeeping attitudes	5 items	.85
LaMonica, et al., 1986 (363)	100,533	IC	Total:	42 items	.92-.95[g]
			Technical/prof.	14 items	.81-.85
			Trust	18 items	.84-.90
			Education	10 items	.80-.84

Abbreviations: AF = alternate-forms reliability estimate; IC = internal-consistency reliability estimate; TRT = test-retest reliability estimate (see text for definitions).

[a] Numbers in parentheses refer to numbered entries in the list of references at the end of this report.

[b] Authors frequently reported reliability for subdimensions and for dimensions that were the sum of two or more subdimensions. To indicate this in the table, the subdimensions included in a dimension are indented and listed immediately after the dimension. NA indicates that publication did not specify number of items used to measure a particular dimension.

[c] Ware and Snyder studied each dimension with more than one measure. The number of measures is shown in parentheses after the dimension name; the number of items per measure is shown as in other table entries.

[d] Breslau and Mortimer reported only the range of coefficients across dimensions (.52 to .82).

[e] Davies, et al., reported a range across four different insurance plan groups.

[f] Wales, et al., reported two TRT coefficients, first for the same interviewer of each administration, second for different interviewers.

[g] LaMonica, et al., reported results from three different studies.

SOURCE: Office of Technology Assessment, 1988.

Only two studies estimated the number of patients required to obtain reliable multi-item scores for individual physicians (180,181). These studies found that 10 patients per physician were inadequate for precise comparisons among individual providers. Having too few patients per provider is a noteworthy shortcoming of the few quality-of-care studies published to date that compare individual physicians. Research in progress suggests that about 40 patients per provider may be required to obtain a reliable quality-of-care score for each provider (683).

VALIDITY OF THE INDICATOR

Of great interest in this evaluation is whether patients' ratings and reports of a given aspect of care reflect known differences in that aspect and not others. Measures used to quantify such differences are referred to as validity variables. In the selection of studies for review, publications were favored that reported analytic methods and findings in sufficient detail to know whether the association between a patient assessment of quality and a validity variable was statistically significant.

Many approaches can be used to evaluate the validity of patients' ratings or reports as an indicator of quality. Regardless of the approach, the purpose is 1) to determine the relationship of patients' assessments to a validity variable (convergent evidence of validity), and 2) to demonstrate that patients' assessments have a weaker relationship with measures of other aspects of quality (discriminant evidence of validity). Because there is no "gold standard," or no one indicator of the "true" quality of medical care, studies pertinent to the validity of patient information about quality of care rely on proxy indicators that vary depending on the quality-related aspect of care being considered. To illustrate, one kind of validity variable would be appropriate for validating patients' assessments of access (e.g., measures of actual office waiting times); another would be appropriate for validating a patient assessment of technical aspects of quality (e.g., independent observation of the process of diagnosis and management). In the OTA literature review, validity standards were applied in a manner consistent with generally accepted guidelines (42).

OTA's review of the considerable literature on patients' assessments of ambulatory care excluded studies that used interventions or measures as validity variables that did not distinguish specific aspects of quality. Because the development and validation of patients' ratings of inpatient care have lagged behind that of ambulatory care, evidence was considered from inpatient studies that linked measures of "overall" quality (e.g., a mix of interpersonal and technical features) to patients' assessments.

Table 11-3 indicates the number of publications included in OTA's indepth review of the evidence on the validity of patients' assessments of quality by setting and aspect of care. Of the 30 publications reviewed, 23 relate to ambulatory care and 7 relate to inpatient care. Nine of the studies manipulated quality-related aspects of care experimentally to test the validity of patients' assessments. Another eight used observational methods to collect data on validity variables (e.g., videotapes of provider-patient encounters) and described elements of the encounters according to objective coding schemes. Seven studies relied on provider report and/or medical chart review to obtain data for validity variables. The remaining studies (as well as some of the preceding ones) used other sources of information (e.g., data on physician/population ratios, staff ratings of ward performance) as validity variables.

The possibility of patient bias affecting the validity of patients' assessments is discussed in the next section. Then, the following sections summarize evidence pertinent to the validity of patients' ratings and reports as indicators of the quality of ambulatory and inpatient care. The evidence is organized within care setting according to the aspect of care described by the validity variable—interpersonal features, technical process, and, for inpatient care only, overall quality.

Table 11-3.—Number of Studies Reviewed by OTA on the Validity of Patients' Assessments of the Quality of Medical Care

Setting, aspects of care, and type of patient assessment	Number of studies[a]
Ambulatory setting:	**23 studies total[a]**
Interpersonal aspects of care	
Patients' ratings	17
Patients' reports	0
Technical aspects of care	
Patients' ratings	9
Patients' reports	3
Inpatient setting:	**7 studies total[a]**
Interpersonal aspects of care	
Patients' ratings	3
Patients' reports	0
Technical aspect of care	
Patients' ratings	2
Patients' reports	3

[a] Because some of these studies covered both interpersonal and technical aspects of care, the figures given below do not add up to the total.

SOURCE: Office of Technology Assessment, 1988.

Bias in Patients' Ratings

Bias in patients' ratings of the quality of medical care has received little empirical attention, and much of the evidence that exists is difficult to interpret. The tendency of people to agree with attitude statements regardless of their content has been shown to bias scores for patient rating instruments when the instruments are not properly balanced (682). Balanced instruments contain both favorably and unfavorably worded statements of opinion about quality. Because they tend to acquiesce, respondents with low socioeconomic status tend to give inflated quality ratings when favorably worded items are relied upon and deflated ratings when unfavorably worded items are relied upon. Thus, in comparisons with scores for more socioeconomically advantaged patient groups, scores for poor patients are biased upward or downward, depending on the type of unbalanced instrument presented to them. Balanced rating instruments have been shown to eliminate this source of bias (682). In studies that rely on unbalanced instruments, however, this source of bias warrants attention when comparisons are made between patient groups differing in socioeconomic status.

Rating bias due to "socially desirable response set," which has been extensively studied in personality research, has also been examined in relation to patients' ratings of quality of care (287). As hypothesized, because it is socially desirable to have a good doctor, patients who present themselves favorably tend to inflate their ratings of the medical care they personally receive. This bias, which tends to be greater among socioeconomically disadvantaged patients, may account, at least in part, for the fact that such patients tend to rate the quality of their care more favorably than their more advantaged patient counterparts. However, the effect of this response set appears to be very weak.

There has been much published discussion of the meaning of significant correlations between patients' sociodemographic characteristics and their quality-of-care ratings (687). Older patients, for example, tend to rate the quality of their care more favorably than younger patients. It is important to keep in mind that these associations tend to be very weak. Further, such findings are difficult to interpret without knowing more about any actual differences in their care. Do older patients rate their care more favorably because they have different preferences or lower standards or because they tend to be treated better? Without independent data about the quality of the care they receive, there is no basis for interpreting correlations between patients' sociodemographic characteristics and their quality-of-care ratings.

Finally, it is sometimes argued that patients' quality-of-care ratings reflect attitudes about life in general (e.g., attitudes toward the community) and are biased by the patients' health status. Findings from a recently completed experiment designed to test for these sources of bias question these arguments (688). Only one of eight correlations between life satisfaction and patients' ratings of their medical care was significant; none accounted for as much as 5 percent of the variance in patients' ratings. All correlations between ratings of personal health status and quality of care were also very weak. Significant correlations were positive, as would be expected, if both health outcomes and patients' quality-of-care ratings are affected favorably by the actual quality of their care. There is no basis for interpreting these results as evidence of bias in patient ratings.

Validity of Patients' Assessments of Ambulatory Care

Interpersonal Aspects of Ambulatory Care

Information from 17 studies that met OTA's review criteria and were relevant to whether data from patients reflect the interpersonal aspects of ambulatory encounters is summarized in table 11-4.[5] Validity variables included experimental manipulation of the provider's behavior toward the patient and independent observation and classification of the provider's affect during an encounter.

The results from the studies shown in table 11-4 indicate that when patients do not rate their providers very favorably in terms of interpersonal manner and skills, in fact providers tend not to be familiar with or knowledgeable about the patient; not to be very skilled in dealing with patient feelings; not to be likely to encourage, support, and involve the patient in care; or not to be courteous, communicative, and relaxed and nonantagonistic in dealing with the patient.

Experimental studies indicate that when interpersonal and technical aspects of the provider's behavior are unrelated during an encounter, patients' ratings accurately distinguish different levels of the two, and their ratings of interpersonal features are not influenced by variations in technical process (150,689).

---

[5] Virtually all of the studies listed in table 11-4 administered satisfaction measures to collect data from patients. Although some items in these measures can be considered reports (e.g., "doctors respect their patients' feelings"), most are evaluative statements (e.g., "doctors always do their best to keep the patient from worrying"). Given the predominance of evaluative items in these studies, they shed most light on the validity of patients' ratings.

Table 11-4.—Validity of Patients' Assessments of the Interpersonal Aspects of Ambulatory Care: Findings From Studies Reviewed by OTA

Study[a]	Sample	Validity variable(s)	Summary of findings
Bertakis, 1977 (66)	100 patients in two studies 1 year apart	Coding of tape-recordings for information given by physician (explanations, tests, regimen, treatment) and amount retained by patient for experimental and control groups	Ratings of interpersonal and technical quality were correlated with the amount of information provided by physician actually retained by patient
Stewart, et al., 1979 (601)	299 visits to 5 physicians	Concordance between physician and patient reports of patients' social problems	Satisfaction ratings 3-mo post-encounter were unrelated to physician knowledge of social problems
Stiles, et al., 1979 (602)	52 patients of 19 physicians in hospital outpatient clinic	Coding of physician behavior in three segments of interview (history, physical, conclusion) in terms of attentiveness, experience, acknowledgment of other's frame of reference, and focus on others	Ratings of interpersonal behavior were positively related to attentiveness in conclusion (r = .43); rating of information-giving was related to informativeness in conclusion, ratings were unrelated to attentiveness during history and physical
DiMatteo, et al., 1980 (181)	462 patients, inpatient and outpatient, of 71 residents in large community hospital	Physician scores on objective measures of ability to interpret affective behavior	Ratings of interpersonal aspects of care delivered by residents were positively correlated with ability to communicate; ratings of technical quality were unrelated to ability
Breslau and Mortimer, 1981 (92)	369 parents of disabled children sampled from 4 clinics	Continuity of care defined in terms of how frequently parent and child saw same physician	Continuity of care was positively related to ratings of interpersonal care, technical quality, finances, and satisfaction in general; highest correlation was with ratings of interpersonal care
Smith, et al., 1981 (583)	29 new patient interviews with 11 physicians	Videotapes of interviews scored for length and interaction process	Ratings (content not given) were more favorable when physicians gave more information, spent more time discussing prevention; unrelated to amount of agreement, casual conversation, suggestions or opinions
Weinberger, et al., 1981 (695)	88 adult outpatient visits with 20 housestaff	Videotaped recordings coded for verbal and nonverbal physician behavior	Patients' satisfaction was higher in encounters with more physician encouragement, coverage of psychosocial issues, and reference to prior visits
Carter, et al., 1982 (123)	101 new patient visits	Trained coders classified physician-patient interaction	Patients' satisfaction was lower when physicians were tense, antagonistic; patient satisfaction was linked to physicians' orientation of patient as to what is being done and why
Comstock, et al., 1982 (144)	10 adult patients for each of 15 residents	Trained observers viewed encounters; rated courtesy and information-giving, coded nonverbal behaviors	Post-visit patient satisfaction ratings (chiefly interpersonal items) were correlated positively with courtesy, attention, listening, empathy, and information-giving
Bartlett, et al., 1984 (56)	60 patients of 5 residents in primary care residency program	Trained observers coded videotaped encounters using interpersonal skills scale	Summary satisfaction score (interpersonal skill, information-sharing, quality of care) was correlated positively (r = .24) with observer ratings of interpersonal skills
Chang, et al., 1984 (128)	268 elderly women volunteers assigned randomly to view simulated encounters	Videotaped patient encounters with nurses and physicians designed to simulate differences in technical quality, patient participation, and handling of psychosocial problems	Overall satisfaction ratings were higher for higher levels of all three manipulations; ratings of visit length, technical quality, and psychosocial care were affected only by technical quality manipulation
Corah, et al., 1984 (150)	24 adult patients of 2 dentists	Experimental manipulation of dentist-patient interaction in terms of amount said, acceptance, reassurance	Ratings of information/communication and understanding/acceptance were more favorable for maximum interaction group; ratings of technical competence were not affected

Table 11-4.—Validity of Patients' Assessments of the Interpersonal Aspects of Ambulatory Care: Findings From Studies Reviewed by OTA (Continued)

Study[a]	Sample	Validity variable(s)	Summary of findings
Stewart, 1984 (600)	140 patients of 24 family physicians	Coders trained in interaction process analysis coded audiotape recordings of physician-patient encounters	Ratings of physician's personal qualities were more favorable 10 days post-visit if encounter had been more patient-centered; ratings of professional competence were higher for physicians showing more tension or asking for opinion/help
Carney and Mitchell, 1986 (121)	120 1st and 3rd year medical students and 60 simulated patients	Faculty assessments of student's overall clinical performance used to form two groups (satisfactory and unsatisfactory)	Ratings of student's interpersonal manner by simulated patients were more favorable for students rated satisfactory by faculty
DiMatteo, et al., 1986 (180)	239 patients of 28 family practice residents in county hospital; outpatient and inpatient	Physician scores on objective measures of ability to interpret affective behavior	Ratings of affective behavior were significantly related (r = .39) to affective ability; ratings of communication and technical quality were unrelated to ability
Cope, et al., 1986 (149)	424 new outpatients of 68 internal medicine residents in large teaching hospital	Evaluations of physician performance by nurses and by supervising faculty	Patients' ratings of physician performance were positively correlated with nurse evaluations (r = .33) and with those by supervising faculty (r = .40)
Ware, et al., forthcoming (689)	109 volunteers randomized to view simulated encounters	Videotaped encounter with physician designed to simulate differences in interpersonal aspects of care (e.g., warmth, communication style)	Interpersonal aspects were rated higher for high than low interpersonal encounters; ratings of technical quality unrelated to interpersonal aspects manipulation

[a] Numbers in parentheses refer to numbered entries in the list of references at the end of this report.

SOURCE: Office of Technology Assessment, 1988.

Technical Aspects of Ambulatory Care

Patients' Ratings.—Entries in the top portion of table 11-5 summarize information about the six studies reviewed by OTA that were relevant to whether patients' ratings reflect the technical process of ambulatory care. Three studies used independent reports or manipulation of number and type of services performed as validity variables (162,377,586). Two studies experimentally manipulated the appropriateness of elements of history-taking and physical examination (128, 689), and one (378) used independent judgments of technical quality.

These studies offer preliminary answers to three questions that have been raised regarding the validity of patients' ratings of technical process: Can consumers distinguish care judged technically good or poor by physicians? Are consumers "seduced" by the kind or number of procedures into believing that services provided were appropriate? Does a provider's interpersonal manner interfere with the patient's accurate assessment of technical process?

Patients' ratings of the technical quality of ambulatory care appear to be somewhat inflated in comparison to ratings made by physicians (689). Despite this, evidence from two experiments in which manipulations of technical process were verified and rated by physicians suggests that, at least for common problems (e.g., chest pain in elderly patients, upper respiratory infection), patients' ratings of the completeness and thoroughness of care accurately distinguish between encounters for which technical process was judged good and less-than-good by physicians (128,689). Another study also found that patients' ratings of the overall quality of care are sensitive to documented variations in technical process (378).

Findings from two studies suggest that patients' ratings of technical process do reflect, at least in part, how many services they received (162,586). However, Linn found no relationship between patients' satisfaction and the number of services they received (377). Results from two experiments (150,689) and two observational studies (180,181) indicate that the physician's interpersonal man-

Table 11-5.—Validity of Patients' Assessments of the Technical Aspects of Ambulatory Care: Findings From Studies Reviewed by OTA

Study[a]	Sample	Validity variable(s)	Summary of findings
PATIENTS' RATINGS			
Linn, 1975 (377)	1,739 encounters in 11 outpatient facilities	Number and type or services performed (e.g., history, exam, lab tests, X-rays)	Summary satisfaction score was unrelated to number or type of services performed
Sox, et al., 1981 (586)	176 outpatients in VA hospital with chest pain	Performance or nonperformance of diagnostic tests	Patients receiving tests rated care for chest pain better than usual, and were less likely to feel that too few tests were done; there were no differences in ratings of interpersonal care and communication
Linn, 1982 (378)	1,418 patients in 20 emergency rooms	Technical process of burn care judged against clinical algorithm	Patients' ratings of overall emergency room care were significantly less favorable with more deviations from algorithm
Chang, et al., 1984 (128)	268 elderly women volunteers assigned randomly to view simulated encounters	Videotaped patient encounters with nurses and physicians designed to simulate differences in technical quality (whether relevant medical history and physical examination items were performed)	Overall satisfaction was rated greater for high than low technical encounters; satisfaction with visit length, technical quality, and psychosocial care was greater for high than low technical encounters
Davies, et al., 1986 (162)	1,537 nonaged adults sampled from general populations	3- to 5-yr followup of groups randomized to HMO or fee-for-service care; expenditures on use 25% lower at HMO	All but low-income, initially well subgroup rated technical quality of fee-for-service care more favorably than HMO care
Ware, et al., forthcoming (689)	109 volunteers randomized to view simulated encounters	Videotaped encounter with physician designed to simulate differences in technical quality (whether relevant medical history and physical examination items were performed)	Technical quality was rated higher for high than low technical encounters; ratings of interpersonal features were unrelated to manipulation of technical quality
PATIENTS' REPORTS			
Cerbert and Hargreaves, 1987 (247)	214 COPD patients of 63 physicians	Physician reports of technical elements of the outpatient visit	There was agreement between physicians' and patients' reports on tests ordered, 96%; treatments mentioned, 94%; occurrence of patient education, 88%
Gerbert, et al., in press (248)	197 COPD patient of 83 physicians	Videotaped outpatient visits checked for mention of theophylline prescription	Patients' reports on interview of having a prescription were in strong agreement with videotaped observations (kappa = 0.05, $p < .001$)
Ware, et al., forthcoming (689)	109 volunteers randomized to view simulated encounters	Videotaped encounter with physician designed to simulate differences in technical quality (whether relevant medical history and physical examination items were performed)	Technical quality was rated higher for high than low technical encounters; ratings of interpersonal features were unrelated to technical quality manipulation

Abbreviations: COPD = chronic obstructive pulmonary disease; HMO = health maintenance organization.

[a] Numbers in parentheses refer to numbered entries in the list of references at the end of this report.

SOURCE: Office of Technology Assessment, 1988.

ner has an insignificant effect on patients' ratings of technical process (see table 11-4).[6]

Patients' Reports.—Entries in the lower portion of table 11-5 summarize information about the three studies OTA reviewed that were relevant to whether patients' reports accurately reflect elements of the technical process of ambulatory care. Results illustrate the relatively high accuracy of patients' reports regarding elements of ambulatory care. Volunteers in the experiments by Ware and colleagues identified medical history and physical examination items that were and were not done with 70 to 88 percent accuracy; better-educated respondents were more accurate (689). In other studies, patients with chronic obstructive pulmonary disease were very accurate (when compared with physicians) in reporting tests ordered (96 percent), treatments mentioned (94 percent), and occurrence of patient education (88 percent) (247), and in reporting prescription medications (when compared with data from videotaped encounters) (248).

---

[6] Results from these four studies were summarized in table 11-4, because the validity variables related to interpersonal aspects of care.

Validity of Patients' Assessments of Inpatient Care

Interpersonal Aspects of Inpatient Care

Entries in table 11-6 summarize information from three studies reviewed by OTA that were relevant to whether data from patients reflect the interpersonal features of inpatient care. All three were experiments; the interventions focused on modifying aspects of provider behavior toward patients by medical staff (374), nursing staff (299), or both (340).

All three studies provided evidence of the convergent validity of patients' assessments. Inpatients' ratings of the interpersonal features of inpatient care that were manipulated experimentally (e.g., communication, involvement in care) were significantly higher for the groups that received the interventions. Ratings of technical/professional aspects of nursing care (299) and of inpatient care overall (340) were also sensitive to these interventions.

Technical Aspects of Inpatient Care

The single entry in table 11-7 summarizes information related to whether patients' assessments are valid reflections of the technical aspects of inpatient care. Because the study by Ehrlich and colleagues listed in table 11-7 is the only study that examined the technical process of inpatient care, the criteria were relaxed somewhat to include it in OTA's literature review. The validity variable used in the study (physicians' judgments of technical process based on medical record review) is not the best standard against which to test patient ratings, given recognized problems with information gaps in medical records. Findings from the Ehrlich study indicate that patients' overall judgments of the quality of medical care delivered during hospital episodes were inflated in comparison to judgments made by physicians, but were more likely to be favorable if care was judged good (as opposed to less-than-good) by the physicians (195).

Table 11-6.—Validity of Patients' Assessments of the Interpersonal Aspects of Inpatient Care: Findings From Studies Reviewed by OTA

Study[a]	Sample	Validity variable(s)	Summary of findings
Ley, et al., 1976 (374) . . .	63 inpatients at hospital in Great Britain	Random groups experiment of extra physician visit to assess, aid patient understanding	Experimental group patients rated communication significantly higher than controls (no visit) or placebo group (visit, no information content)
Hinshaw, et al., 1983 (299)	88 surgical patients	Random groups experiment of perioperative registered nurse visits to reassure and educate; quality independently judged better for visited patients	Patients' ratings of trusting relationship and technical/professional nursing care were significantly more favorable for visited group
Kane, et al., 1985 (340) . .	246 inpatients in VA hospital	Random groups experiment of hospice ward/team intervention that increased provider communication, more patient/family involvement in care[b]	Patients' ratings of involvement in care and care overall (technical, interpersonal, general) were significantly higher for hospice group; ratings of physical environment were unaffected by intervention

[a]Numbers in parentheses refer to numbered entries in the list of references at the end of this report.
[b]Manipulation check not reported by authors.

SOURCE: Office of Technology Assessment, 1988.

Table 11-7.—Validity of Patients' Assessments of the Technical Aspects of Inpatient Care: Findings From the Study Reviewed by OTA

Study[a]	Sample	Validity variable(s)	Summary of findings
Ehrlich, et al., 1961 (195) . .	283 Teamsters in 105 New York hospitals	Physician judgments of technical quality based on record review	Patients' ratings of medical care while hospitalized were significantly related to physician judgments of technical quality: 5 of 6 stays rated not good by patients were judged fair or poor by physicians; significantly more patients judged care best when rated excellent or good by physicians (86% vs. 74%)

[a] Numbers in parentheses refer to numbered entries in the list of references at the end of this report.

SOURCE: Office of Technology Assessment, 1988.

Overall Quality of Inpatient Care

Patients' Ratings.—Entries in the top portion of table 11-8 summarize information from two studies reviewed by OTA that were relevant to whether patients' ratings reflect the overall quality of inpatient care. Validity variables included summary rankings of psychiatric wards by staff on a range of criteria (517) and recommendations for care made by nurses (299).

Results from both studies support the validity of patients' ratings of the overall quality of inpatient care. Rice and colleagues noted that rankings of psychiatric wards from patients' overall ratings were identical to the rankings made by staff (517). In the study by Hinshaw and colleagues, patients' ratings of overall quality were significantly higher when nurses made more recommendations regarding care; researchers presumed that more recommendations reflected better quality nursing care (299).

Patients' Reports.—Entries in the bottom portion of table 11-8 summarize information from the three studies included in OTA's review that were at all relevant to whether patients' reports reflect the overall quality of inpatient care. Validity variables included staff reports of omissions in nursing care (1), staffing levels of professional nurses (2), and reviews of patients' medical records (195). Results provide an equivocal answer to the question of whether patients' reports are sensitive to the overall quality of inpatient care, in part because none of the studies had well-defined validity variables and in part because two (1,195) of the three reported results in such a way that true rates of underreporting (or overreporting) could not be discerned.

Abdellah and Levine reported that 100 percent of omissions in nursing care (e.g., failure to administer medications on schedule, failure to answer call bell) reported by inpatients were verified by staff (1). Because staff were asked to verify only those omissions reported by patients, one cannot be sure that underreporting of omissions did not occur.

A later study by Abdellah and Levine demonstrates the sensitivity of inpatients' reports about the quality of nursing care to staffing levels of registered nurses (2). Patients reported fewer omissions in care for which registered nurses would be expected to be more responsible (e.g., therapy) when there were relatively more registered nurses. By contrast, patients' reports about things for which registered nurses were not primarily responsible (e.g., attention to dietary needs) were unrelated to professional/nonprofessional nurse staffing ratios.

Ehrlich and colleagues found that a substantial minority (one-third) of patients underreported the diagnostic tests they had prior to a hospitalization (195). Given the way the authors reported their data, one cannot determine from this study the number or type of tests that were underreported or the effect of the timing of the patient survey.

Table 11-8.—Validity of Patients' Assessments of the Overall Quality of Inpatient Care: Findings From Studies Reviewed by OTA

Study[a]	Sample	Validity variable(s)	Summary of findings
PATIENTS' RATINGS			
Rice, et al., 1963 (517)	457 psychiatric inpatients	Sum of staff rankings of ward on seven criteria[b]	Patients' rankings of overall hospital care (sum of physical facilities, patient services, and patient management) were identical to staff rankings
Hinshaw, et al., 1983 (299)	88 surgical patients	Number of care recommendations made by registered nurses (more presumed to indicate better quality nursing care)	Patients' ratings of trust, technical, education, and overall hospital care were significantly more favorable when more recommendations made
PATIENTS' REPORTS			
Abdellah and Levine, 1957 (1)	60 inpatients at a Midwestern hospital	Query of staff member to determine whether reported omission in nursing care had occurred	100% of patient-reported omissions in nursing care verified by staff report
Abdellah and Levine, 1958 (2)	9,000 inpatients in 60 Midwestern hospitals	Staffing levels of registered nurses (higher levels presumed to indicate better quality of care)	There were significantly fewer patient-reported omissions in nursing therapy with more professional registered nurses; staffing levels were weakly or not at all related to reported omissions in environmental features or attention to dietary needs
Ehrlich, et al., 1961 (195)	283 Teamsters in 105 New York hospitals	Chart review to identify diagnostic tests prior to a hospitalization	Tendency to underreport: 1/3 of patients failed to report tests mentioned in chart

[a] Numbers in parentheses refer to numbered entries in the list of references at the end of this report.

[b] Criteria included adequacy of physical facility; crowdedness; patient morale; staff morale; amount of staff/patient contact; degree of patient/staff harmony; and amount of freedom granted patients.

SOURCE: Office of Technology Assessment, 1988.

FEASIBILITY OF USING THE INDICATOR

There are three basic questions regarding the feasibility of obtaining data from patients on quality-related attributes of care:

- Are appropriate survey instruments and data collection techniques available and/or can they be developed?
- Can potential respondents to patient surveys who will agree to respond be identified?
- Are the costs of obtaining data from consumers reasonable?

Each can be answered affirmatively. In part because the literature reviewed in this chapter directly addresses only the second question (by reporting response rates), the answers depend heavily on practical experiences and knowledge of the literature on survey research methods in general, a detailed synthesis of which was beyond the scope of this review.

There are several good survey instruments that can be used to obtain patients' ratings of ambulatory care, "good" meaning that the instruments do a comprehensive job of representing one or more attributes of care for which patients provide valid data (for examples, see the studies cited in tables 11-4 and 11-5). Published instruments for obtaining patients' ratings of inpatient care rarely have done a good job of this. Part of the reason is that there is considerably less information about the dimensions of hospital care that relate to quality than about the dimensions of ambulatory care (544) and that fewer studies have examined the validity of published instruments pertaining to inpatient care. An ongoing collaborative effort by the Hospital Corporation of America, Harvard Community Health Plan, and the Rand Corporation to develop and test a hospital satisfaction survey should provide useful information in this regard (69).

Photo credit: Metropolitan Health Services Center

Standardized survey instruments for collecting valid patients' ratings, particularly for the inpatient setting, have not been developed.

A wide variety of techniques for collecting data are available and have been used to obtain information from consumers. Self- and interviewer-administration of survey instruments (usually in person; sometimes by telephone) are the most commonly used. The "best" technique will vary depending on the study population, the complexity of the data collection instrument, and a variety of other considerations.

A ready way of identifying potential respondents for patient surveys would be through the management information systems available in many ambulatory and inpatient settings. Depending on the focus of a particular quality of care evaluation, management information systems could identify, for example, a universe of patients (in the case of an enrolled population): users versus nonusers; patients who complain or lodge formal grievances; hospitalized patients, by admitting diagnosis, procedure, or unit; and patients who see a particular provider.

Because patients are generally willing to discuss their medical care experiences and attitudes, good response rates (70 percent or higher) can be achieved on patient surveys (3,69,102,235,691). Lower response rates, which raise questions about sample bias, are often caused by inadequate followup efforts.

Few published studies include any information about how the costs of collecting data on the quality of medical care from patients compare with costs of obtaining data from other, more traditional sources (e.g., medical record audit, computerized claims audit). Survey costs will vary markedly depending on such factors as administration method, dispersion of the sample, availability of potential respondents, followup procedures, and questionnaire length. Mail and telephone surveys usually cost considerably less than personal interviews (225).

What little evidence is available suggests that information acquired from patients costs no more, and in many circumstances less, than information obtained from medical record reviews. Recently obtained cost estimates suggest that medical record abstractions designed for evaluating the quality of care range from $35 to $45 per record (161); costs for mail and telephone surveys of typical length (15 to 20 minutes) appear to range from about $15 to $45 dollars (69,161,680). Of course, data obtained for quality assessment as a byproduct of existing data systems, such as hospital discharge abstracts or billing claims, would be much less costly.

Whether obtaining information from patients is a cost-effective method of obtaining data on indicators of the quality of care is an open question. Only one study identified in OTA's literature review compared the costs of obtaining quality-relevant data from different sources (physician and patient interviews, medical record abstracts, and coding of videotaped encounters) (247). Findings from that study illustrate forcefully that data from all sources on the technical aspects of care should be used as complements, rather than as substitutes, until research can better identify which source provides the most accurate (and least expensive) information (247, 248). Given the paucity of data from other traditional sources of information regarding the interpersonal aspects of medical care, and the intrusiveness, complexity, and cost of using approaches such as direct observation and coding of the provider-patient encounter (316), obtaining information from patients appears to be the most cost-effective approach for assessing the interpersonal aspects of the quality of care.

CONCLUSIONS AND POLICY IMPLICATIONS

On the basis of the review in this chapter, one may conclude that it is possible to construct valid patient-based indicators of the quality of medical care and that there are good reasons to use such indicators given the shortcomings inherent in alternative strategies. By all standards considered, there is a strong case for using patients' assessments as indicators of the quality of the interpersonal aspects of care both of physicians in ambulatory settings and of physicians and hospital staff in inpatient settings.

This conclusion about patients' assessments is based on several considerations. On the one hand, there is no practical or valid alternative source of information on the interpersonal manner of physicians and other health care providers described in the literature. Direct observation must be eliminated on grounds of impracticality (intrusiveness, complexity of coding schemes, and expense) and because of concerns about whether ratings by trained observers adequately reflect patients' values. Furthermore, there is no evidence that patients' medical records, in either ambulatory or inpatient settings, provide valid information about the interpersonal aspects of care. Even if providers routinely made notes about the quality of their interpersonal relationships with patients, there would still be reason to question the validity of the notes. Who is more qualified than the individual patient to judge the interpersonal manner of physicians and other health care providers in light of patients' standards?

These arguments themselves, however, provide no guarantee that patients' assessments are valid indicators of the interpersonal quality of care. The crucial pieces of the puzzle are published findings regarding the empirical validity of patients' ratings of the interpersonal aspects of care. OTA's literature review identified considerable evidence that patients' ratings are valid indicators of interpersonal aspects of care in ambulatory and inpatient settings. The evidence across settings comes from 20 studies that compared results from objective measures of the interpersonal aspects of care with patients' ratings. The validation standards in these studies included direct observation by trained observers, evaluations by physicians and other health care providers, analyses of audiotape and videotape recordings, randomized-group experiments to evaluate interventions designed to change the interpersonal aspects of care, and studies of randomized groups in which variations in interpersonal aspects of care were experimentally manipulated. Of the 23 studies of patients' assessments in ambulatory settings that satisfied OTA's selection criteria, 17 yielded evidence in support of the validity of patients' assessments of the interpersonal aspects of quality. Of the 8 studies of patients' assessments in inpatient settings, 3 yielded evidence in support of the validity of patients' assessments of the interpersonal aspects of quality.

Relatively little published evidence was found regarding the validity of patients' assessments of the technical aspects of quality. This dearth of evidence is unfortunate, because most other methods for assessing technical quality, such as medical record audit, carry high dollar and time costs. Further, ambulatory care records (as opposed to hospital inpatient records) are an incomplete source of information about the quality of the technical process of care. The search for data sources that are less costly and that help to fill the gaps leads some to consider surveying patients about their care.

Photo credit: Harvard Community Health Plan

Patients' assessments of the interpersonal aspects of ambulatory care are valid indicators of the quality of care.

The available evidence, although limited and only from ambulatory settings, generally supports the use of patients' ratings of the technical aspects of care as indicators of quality. Specifically, the few available studies that have verified differences in technical process (e.g., physician/staff assessments, experimental manipulations) and have compared results with patients' ratings have consistently linked the two. Moreover, evidence from experimental studies suggests that, at least for relatively common ambulatory conditions, a physician's interpersonal manner does not obscure patients' ability to detect variations in technical process. Nevertheless, pending further research on this issue and replication of these findings, patients' ratings of the technical aspects of care perhaps should be used only in conjunction with highly credible data about the technical aspects for purposes of evaluating the quality of medical care.

A promising but rarely employed strategy for patient-based assessments of the quality of care would be one based on patients' reports of what does and does not occur. This approach makes no assumption about patients' qualifications as judges, only about the accuracy of their reports. Physicians or others using algorithms for evaluating the technical aspects of care can use such patients' reports to make the actual judgments regarding quality of the technical process. Further research is needed to determine what aspects of the technical process can or cannot be reported accurately by patients in order to test the suitability of this strategy in both ambulatory and inpatient settings.

Not surprisingly, available evidence establishes a direct link between the specificity of the content of patients' quality assessments and the validity of such assessments. Technical and interpersonal aspects of care are distinct quality-related attributes that can be measured and interpreted separately. Validity is generally better when there is a good match between the content of the assessment and the quality aspect of interest. More global measures (e.g., overall satisfaction ratings, whether patients are willing to recommend a hospital to others, health care plan disenrollment rates), however, are not unrelated to quality of care. Given the overriding importance of quality of care to consumers, large differences in such global indicators of satisfaction are likely to reflect differences in quality. Because global indicators are sensitive to a wide range of influences, however, other interpretations of such indicators should be kept in mind. Further, global measures are not as programmatically useful, because they do not provide clues as to which aspects of quality are most likely to account for any differences that are observed.

Priorities for future research in the inpatient setting should include studies of patients' assessments of specific features of quality of care that have not been analyzed in work to date, including interpersonal and technical aspects of medical and nursing care, information-giving and other aspects of communication, and patient and family involvement in care. Little is known about how differences in the quality of the inpatient technical process are experienced by patients or how the differences affect patients' assessments. OTA's review has yielded no support, however, for the common belief that patients' assessments of the quality of hospital care are determined by amenities.

To evaluate the validity of patients' assessments, OTA examined the content of published survey instruments to determine how well the instruments reflected patients' values, and their comprehensiveness in relation to the universe of patient experiences. Although a number of published instruments are quite comprehensive, none covers all aspects of quality well. Available taxonomies of patient experiences with ambulatory (687) and inpatient care (544) should be used as minimum standards for judging the content of candidate measures. Published instruments designed to obtain data from patients about hospital care are particularly lacking in this regard, and further developmental work is required to develop useful instruments.

It is likely that quality considerations will be increasingly emphasized in attempts to market prepaid and other group plans, health insurance benefits, and hospital facilities to consumers. Such efforts appeal directly to consumers' desires for good quality health care. This marketing trend underscores both the potential value of published patient-based information regarding the quality of physician and hospital performance and the potential for abuse of the data.

Because of the importance of measurement and patient sampling methods in determining results, there is a need to standardize methods and to develop minimum standards for reporting results to the public. To be valid, comparisons among physicians or hospitals must be based on standardized survey instruments, data collection procedures (e.g., personal or telephone interview, self-administered questionnaire), and survey methods (e.g., timing of administration), as well as on representative samples. Reproducible scores can be achieved only if methods are carefully standardized.

Finally, it can be argued that routine and careful monitoring of patient-based indicators of the quality of physician and hospital care is important regardless of conclusions about the validity of these indicators in measuring true quality. Instead, the argument is based on strong empirical evidence that patients' perceptions of quality of care influence patients' behavior (406,685). Patient behaviors that are affected include doctor-shopping, complaints, disenrollment, compliance, and use of services. Such behaviors have noteworthy consequences to their health and the quality of their care.

Appendixes

Appendix A

Method of the Study

This assessment was prompted by congressional interest in whether valid information on hospital and physician quality could be developed and distributed to the public to assist their choice of health care providers. The study was requested by the House Committee on Energy and Commerce, and endorsed by the Senate Committee on Finance, the Senate Special Committee on Aging, the Subcommittee on Consumer of the Senate Committee on Commerce, Science, and Transportation, and the House Committee on Science, Space, and Technology. The interest of the committees was primarily in measures of quality that could be applied to acute care hospitals and physicians, but the committees were also interested in evaluating the quality of health plans. On September 23, 1987, the OTA project Technology Assessment Board approved the proposal for this project.

During the early part of the project, OTA staff consulted with consumer organizations, professional organizations, unions, employers' associations, third-party payers, health services researchers, and methodologists for suggestions of candidates for the study's advisory panel. The advisory panels for OTA studies guide OTA staff in selecting material and issues to consider and review the written work of the staff, but the panels are not responsible for the content of final reports. The advisory panel for this study consisted of 21 members from parties with expertise or an important perspective: consumer advocacy, medical practice, nursing, hospital management, health insurance, rural health, corporate health benefits, unions, law, health maintenance organizations, quality assessment organizations, State health departments, quality assessment research, information dissemination, and health policy analysis. Frederick Mosteller from the Department of Health Policy and Management at the Harvard School of Public Health chaired the advisory panel for the study

The first meeting of the advisory panel was held on February 3, 1987. Before the meeting, the OTA project staff began preliminary research into the issues involved in selecting and evaluating indicators for quality assessment and prepared a draft outline for the study. During the meeting, panel members were asked to discuss a framework for consumers to assess the quality of care and methods of presenting quality information to consumers. In addition, the panel members discussed the relevant issues relating to quality assessment so as to narrow the scope of OTA's task. As a result of the panel meeting and discussions with congressional staff, the scope of the study was limited to physicians and hospitals.

On March 3, 1987, a workshop was held to consider the procedure that OTA should use to evaluate the reliability, validity, and feasibility of the selected indicators of the quality of medical care. The workshop, chaired by Frederick Mosteller, included members experienced in evaluative research methods (see app. B). On the basis of the comments received from this workshop, the OTA staff revised the evaluation procedure to give more emphasis to measurement issues and developed a checklist to apply to specific studies.

An additional workshop was held on March 23, 1987, for the purpose of developing a list of quality indicators to evaluate for the OTA study and to discuss further the framework to assess quality from a consumer's perspective. This workshop was chaired by R. Heather Palmer, a member of the advisory panel, and included several other panel members (see app. B for a complete list of workshop participants). After this meeting, the OTA staff selected the following eight indicators of quality for evaluation: 1) hospital mortality rates; 2) adverse events that affect patients; 3) formal State disciplinary actions, sanctions recommended by peer review organizations and imposed by the Department of Health and Human Services, and malpractice compensation; 4) the evaluation of physicians' performance as exemplified by care for hypertension; 5) volume of procedures performed by hospitals and physicians; 6) scope of hospital services, with emphasis on emergency services, cancer care, and neonatal intensive care units; 7) physician specialization; and 8) patients' assessments of their care. Also on the basis of the workshop discussion, OTA staff decided to limit the aspects of access to be considered in the report to those that overlapped with quality and pertained once a person had decided to seek care.

Using a method of evaluation developed for this study (see app. C), OTA staff began to evaluate six of the eight indicators selected for evaluation. Contractors were chosen to evaluate the two remaining indicators: volume of procedures performed by hospitals and physicians and patients' assessments of their care.

As OTA staff began to consider the policy implications of the study's findings, it became apparent that they needed further information on certain specialized topics. During the summer of 1987, OTA contracts were let to fill gaps related to the availability of data, legal issues surrounding peer review, the use of quality-of-care information by consumers, organizational loci for constructing and evaluating quality indicators, the validity of malpractice profiles, and legal issues regarding confidentiality of data on physicians (see list below).

The second meeting of the advisory panel was held on July 26-27, 1987, to bring the panel members up to date on the progress of the study and to review preliminary drafts of some sections of the report. OTA staff developed brief descriptions of each indicator for the panel's discussion. The panel gave advice on how to disseminate information on the quality indicators to the public.

During the rest of the summer and fall of 1987, OTA project staff reviewed the literature on the various indicators and compiled the respective evaluations. Throughout this time, draft papers were received from contractors. On the basis of comments from the OTA project staff, advisory panel members, and outside reviewers with expertise in the relevant fields, the contractors revised their papers.

In mid-January 1988, the draft report for the overall study was sent for review to the advisory panel and to a wide range of other experts and interested parties. Discussion of the draft report formed the subject of the final meeting of the advisory panel on February 2-3, 1988. During February and March 1988, the OTA staff revised the report in response to discussion at the final panel meeting and ouside reviewers' comments. The staff prepared a final draft, which was submitted in late March 1988 to the Technology Assessment Board for its approval.

In addition to the main report, other documents prepared to provide background information are available through OTA in limited quantities. Some of these stem from contractors' reports, and others present detailed technical information on specific indicators analyzed by OTA staff.

- Nancy E. Cahill, "Developing Law on Professional Standards and Peer Review in Quality Assessment Activities," Duke University, 1987;
- Denise Dougherty, "Hospital Mortality Rates as a Quality Indicator," Office of Technology Assessment, 1988;
- Karen Glanz and Joel Rudd, "Effects of Quality of Care Information on Consumer Choice of Physicians and Hospitals," University of Minnesota and University of Arizona, 1987;
- Peter G. Goldschmidt, "The Appropriate Organizational Loci for Constructing Indicators of the Quality of Hospitals and Physicians and for Evaluating the Validity of Those Indicators," World Development Group, Inc., 1987;
- Marlene Larks, "Access to Health Data by State Health Data Organizations and Quality Assessors," National Association of Health Data Organizations, 1987;
- Harold S. Luft, Deborah W. Garnick, David Mark, Stephen J. McPhee, and Janice Tetreault, "Evaluating Research on the Use of Volume of Services Performed in Hospitals as an Indicator of Quality," University of California, San Francisco, 1987;
- Mark McClellan, "Hypertension Screening and Management as an Indicator of Quality: An Evaluation of the Literature," Massachusetts Institute of Technology, 1988;
- Don Harper Mills and Orley Lindgren, "Physician Malpractice Profiles as Indicators of Quality: Reliability, Validity, and Feasibility Issues," Institute for Medical Risk Studies, 1987;
- Beth Mitchner, "Physician Specialization as an Indicator of Quality: An Evaluation of the Literature," Office of Technology Assessment, 1988;
- James B. Simpson, "Release of Physician-Specific Quality of Care Information: Legal Issues," Western Consortium for the Health Professions, 1987;
- SysteMetrics, "Report on Available State-Specific Data Bases," 1987; and
- John E. Ware, Jr., Allyson Ross Davies, and Haya H. Rubin, "The Suitability of Consumers' Assessments of Physician and Hospital Performance as Indicators of the Quality of Care," The Rand Corporation, 1987.

Appendix B

Acknowledgments

This project has benefited from the advice and review of several people in addition to the advisory panel. OTA staff would like to express its appreciation to the following people for their valuable guidance.

Betty Jane Anderson
American Medical Association,
Chicago, IL

John T. Ashley and Staff
University of Virginia Hospital
Charlottesville, VA

Jack Beirig
Sidley & Austin
Chicago, IL

Deborah S. Berkowitz
Department of Quality Assurance
Blue Cross and Blue Shield Association
Chicago, IL

Jill Bernstein
General Accounting Office
U.S. Congress
Washington, DC

Mark Blumberg
Kaiser Foundation Health Plan, Inc.,
Oakland, CA

Fred Bodendorf
Pennsylvania Health Care Cost Containment Council
Harrisburg, PA

Patricia Booth
Office of Medical Review
Health Care Financing Administration
Baltimore, MD

Alexander E.M. Borgiel
Practice Assessment Committee
College of Family Physicians of Canada
Mississauga, Ontario

Randall Bovbjerg
Urban Institute
Washington, DC

Dale Breadon
Federation of State Medical Boards
Ft. Worth, TX

John P. Bunker
Division of Health Services Research
Stanford University School of Medicine
Stanford, CA

Howard Champion
Medstar
Washington Hospital Center
Washington, DC

Paul Cleary
Beth Israel Ambulatory Care Center
Boston, MA

Danner Clouser
Milton S. Hershey Medical Center
Pennsylvania State University
Hershey, PA

Francis E. Conrad
Office of Quality Assurance
Veterans Administration
Washington, DC

Patricia Danzon
Wharton School of Business
University of Pennsylvania
Philadelphia, PA

Feather Davis
Office of Research
Health Care Financing Administration
Baltimore, MD

Linda Demlo
General Accounting Office
U.S. Congress
Washington, DC

Susan DesHarnais
Commission on Professional and Hospital Activities
Ann Arbor, MI

Paul Eggers
Office of Research
Health Care Financing Administration
Baltimore, MD

John Finnegan
School of Public Health
University of Minnesota
Minneapolis, MN

Ann Barry Flood
College of Medicine
University of Illinois
Urbana, IL

Jinnet Fowles
Park Nicollet Medical Foundation
Minneapolis, MN

Bryan Galusha
Federation of State Medical Boards
Ft. Worth, TX

Irene Gibson
Health Standards and Quality Bureau
Health Care Financing Administration
Baltimore, MD

Dennis Gold
Professional Affairs and Quality Assurance
Department of Defense
Washington, DC

Willis Goldbeck
Washington Business Group on Health
Washington, DC

Robert Goldenberg
Department of Obstetrics/Gynecology
University of Alabama
Birmington, AL

Marion Gornick
Office of Research
Health Care Financing Administration
Baltimore, MD

Frank Grad
Columbia Law School
New York, NY

Bradford Gray
Institute of Medicine
National Academy of Sciences
Washington, DC

Judith Hall
Department of Psychology
Northeastern University
Boston, MA

Edward Hannan
Bureau of Health Care Research and Information Services
State of New York Department of Health
Albany, NY

Martin J. Hatlie
American Medical Association
Chicago, IL

Kathy Headon
Office of Medical Review
Health Care Financing Administration
Baltimore, MD

Howard Hiatt
Brigham and Women's Hospital
Boston, MA

Judith Hibbard
Department of Health Education
University of Oregon
Eugene, OR

James Hughes
Hospital Infections Program
Centers for Disease Control
Atlanta, GA

Robert Hughes
School of Health Administration and Policy
Arizona State University
Tempe, AZ

Lisa Iezzoni
Health Care Research Unit
Boston University Medical Center
Boston, MA

Steven Jencks
Office of Research
Health Care Financing Administration
Baltimore, MD

Alan Kaplan
American Association of Retired Persons
Washington, DC

Joyce Kelly
Division of Intramural Research
National Center for Health Services Research
Rockville, MD

A.B. Kiesewetter
Acute Care Hospital Regulation
State of California Department of Health Services
Sacramento, CA

Susan Kladiva
General Accounting Office
U.S. Congress
Washington, DC

Henry Krakauer
Health Standards and Quality Bureau
Health Care Financing Administration
Baltimore, MD

Donald G. Langsley
American Board of Medical Specialties
Evanston, IL

Steven LaTour
Marketing Department
Northwestern University
Chicago, IL

Bruce Lehman
Swindler & Berlin
Washington, DC

Arthur Levin
Center for the Medical Consumer
New York, NY

William Libercci
Office of the Inspector General
U.S. Department of Health and Human Services
Baltimore, MD

Richard Lichtenstein
Health Services Management and Policy
School of Public Health
University of Michigan
Ann Arbor, MI

Karen S. Liebert
University of Virginia Hospital
Charlottesville, VA

William Lohr
Division of Extramural Research
National Center for Health Services Research
Rockville, MD

Lisa Looper
American Medical Peer Review Association
Washington, DC

James Lubitz
Office of Research
Health Care Financing Administration
Washington, DC

Karl Mangold
The Fisher-Mangold Group
Pleasanton, CA

James Maroc
Iowa Foundation for Medical Care
West Des Moines, IA

William McAuliffe
Department of Psychiatry
Cambridge Hospital
Cambridge, MA

Daisy McGinley
General Accounting Office
U.S. Congress
Washington, DC

Bruce McPherson
American Hospital Association
Chicago, IL

Judith Moore
Prospective Payment Assessment Commission
Washington, DC

Michael Moran
Health Standards and Quality Bureau
Health Care Financing Administration
Baltimore, MD

Laura Morlock
Health Services Research and Development Center
John Hopkins University
Baltimore, MD

Eugene C. Nelson
Department of Community and Family Medicine
Dartmouth Medical School
Hanover, NH

Duncan Neuhauser
Department of Community Health
Case Western Reserve University
Cleveland, OH

Arthur Osteen
American Medical Association
Chicago, IL

Mary A. Philipsen
Office of Medical Review
Health Care Financing Administration
Baltimore, MD

Susan Polniaszek
United Seniors Health Cooperative
Washington, DC

Gerald Riley
Office of Research
Health Care Financing Administration
Baltimore, MD

Barbara Rimer
Behavioral Research
Fox Chase Cancer Center
Philadelphia, PA

Shirlee Rivers
American Medical Association
Chicago, IL

Leslie Roos
Department of Business Administration
The University of Manitoba
Winnepeg, Manitoba

Dale Rublee
American Medical Association
Chicago, IL

Michael J. Saks
College of Law
University of Iowa
Iowa City, IA

Marcel Salive
Public Citizen Health Research Group
Washington, DC

George Schieber
Office of Research
Health Care Financing Administration
Baltimore, MD

Ronald Schoenberg
Laboratory for Social and Environmental Studies
National Institute of Mental Health
Bethesda, MD

Lynn Silver
Public Citizen Health Research Group
Washington, DC

Frank A. Sloan
Department of Economics
Vanderbilt University
Nashville, TN

John Spiegel
Health Standards and Quality Bureau
Health Care Financing Administration
Washington, DC

Debra Stashower
Trial Lawyers Association
Washington, DC

Robin Strongin
Leadership Commission on Health Care
Washington, DC

Steven J. Summer
Maryland Hospital Association
Lutherville, MD

Beth Tandski
Division of Medicine
Health Resources and Services Administration
Rockville, MD

Emilio Venezian
Department of Business Administration
Rutgers University
Newark, NJ

Douglas P. Wagner
ICU Research
George Washington University Medical Center
Washington, DC

Andrew Watry
Composite State Board of Medical Examiners
Atlanta, GA

Neil Weisfeld
Licensure Reform Project
New Jersey Department of Health
Trenton, NJ

Norman Weissman
Division of Extramural Research
National Center for Health Services Research
Rockville, MD

Margaret Wilson
Division of Medicine
Health Resources and Services Administration
Rockville, MD

Marian Wiseman
American College of Emergency Physicians
Dallas, TX

Mark Yessian
Office of Inspector General
U.S. Department of Health and Human Services
Boston, MA

WORKSHOP ON THE SYSTEMATIC EVALUATION OF AVAILABLE INDICATORS, MAR. 3,1987

Deborah Garnick
Institute for Health Policy Studies
University of California
San Francisco, CA

Laura Leviton
Department of Health Services Administration
Graduate School of Public Health
University of Pittsburgh
Pittsburgh, PA

Kathleen Lohr
Council on Health Care Technology
Institute of Medicine
Washington, DC

William McAuliffe
Department of Psychiatry
Cambridge Hospital
Cambridge, MA

Frederick Mosteller
Department of Health Policy and Management
Harvard School of Public Health
Boston, MA

Leonard Saxe
Center for Applied Social Sciences
Boston University
Boston, MA

Ronald Schoenberg
Laboratory for Social and Environmental Studies
National Institute of Mental Health
Bethesda, MD

Frederic Wolf
Department of Postgraduate Medicine and Health Professions Education
University of Michigan
Ann Arbor, MI

Paul Wortman
Wortman & Associates
Ann Arbor, MI

WORKSHOP ON THE FRAMEWORK TO ASSESS THE QUALITY OF MEDICAL CARE FROM A CONSUMER PERSPECTIVE AND SELECTION OF INDICATORS FOR OTA EVALUATION, MAR. 23, 1987

Robert Brook
The Rand Corporation
Santa Monica, CA

Avedis Donabedian
School of Public Health
University of Michigan
Ann Arbor, MI

Kathleen Lohr
Council on Health Care Technology
Institute of Medicine
Washington, DC

Lorna McBarnette
New York State Department of Health
Albany, NY

R. Heather Palmer
Harvard School of Public Health and Institute for Health Research
Boston, MA

Laurence R. Tancredi
University of Texas Health Sciences Center at Houston
Houston, TX

Appendix C

Method Used by OTA To Evaluate Indicators of Quality

Introduction

As part of its assessment, OTA developed a systematic method for synthesizing available information on potential indicators of the quality of medical care. The method OTA developed was oriented to evaluating the reliability, validity, and feasibility of quality indicators generically—that is, it was intended to apply to all quality indicators however measured. OTA developed the method with the assistance of a workshop of experts, including several members of the advisory panel for the entire study (see apps. A and B). OTA used the method to evaluate the quality indicators it selected for intensive review in this assessment. This appendix describes the rationale for employing a systematic method for evaluation, the method OTA developed, and that method's limitations.

Rationale for a Systematic Method

Numerous observers have remarked on the need for systematic syntheses of bodies of scientific literature, as opposed to the more typical "narrative" or "casual" reviews (148,254,291,311,376,489,539,710). Typical narrative reviews have a number of problems (710). Reviewers may include studies selectively or haphazardly rather than surveying systematically the literature base. They may weight studies differently when interpreting a set of findings, for example, giving more credence to studies conducted by widely known authorities, or to studies that appear to have better designs. These two factors can result in misleading interpretations of study findings. Even if the overall interpretation of a set of findings is accurate, reviewers may fail to examine characteristics of the studies as potential explanations for disparate or inconsistent results across studies. Finally, an overall result may hold only in specific circumstances; the casual review may fail to examine moderating variables.

As a result of the selective inclusion of studies and differential subjective weighting of studies in the interpretation of a set of findings, conclusions of typical narrative reviews are not able to be compared to one another, even when the reviews address the same topic. OTA planned to evaluate the reliability, validity, and feasibility of a number of indicators, and wished to be able to have the same level of confidence in each evaluation and to make the evaluations themselves readily evaluable. As pointed out by Wolf, it has been argued that the same scientific rigor be applied to research literature reviews as to the individual studies addressing the research question at hand (710).

Description of OTA's Method: Procedure and Checklist for Evaluation

The method OTA developed to evaluate indicators of the quality of medical care actually consists of two parts. The first part, an overall guide to evaluating an indicator, was called the procedure. The second part was called the checklist. Each of these is described below. For more information, see the detailed outline of the procedure and annotated checklist at the end of this appendix.

Procedure for Evaluating an Indicator

The procedure outlined the steps OTA wished all readers[1] to take so that the evaluation of indicators would be as consistent and rigorous as possible, given OTA's resource limitations. These steps included:

- describing the indicator;
- selecting information to evaluate the indicator;
- evaluating the citations selected, including applying and refining the checklist; and
- presenting the method and findings in written form (see attached procedure and checklist).

Particular attention was paid to the method by which citations (e.g., articles, reports of studies) were identified and selected for evaluation, because, as noted above, selective inclusion and exclusion of studies are potential sources of bias in literature reviews.

Most research syntheses are based exclusively on published studies from the scientific literature. OTA found, however, that for some indicators, such as disciplinary actions, there were few or no published studies. In such cases, OTA relied on other sources of information, such as descriptions of procedures of State medical boards. In addition, much of OTA's evaluations of feasibility relied on the staff's general knowledge of the health care system. The factors on the checklist were applied to these other sources of information as well, to evaluate reliability, validity, and feasibility at the indicator level. Thus, the checklist was applied both to particular sources of information and at the indicator level.

---

[1] In this report, OTA staff and contractors who read and evaluated studies pertaining to indicators are referred to as readers to distinguish them from outside reviewers of OTA's work.

OTA staff were trained (in-house) in the use of the Medline and Healthline data bases. All readers, OTA staff as well as contractors, were instructed to maintain good records of all citations considered for evaluation. The procedure called for readers to be trained in use of the checklist as well. OTA staff met several times to clarify items on the checklist, discuss its use, refine it through consensus, and otherwise ensure that it was being applied reliably. Major refinements were to be communicated to contractor readers. As the final step in the evaluation process, the written summaries of the evaluations were reviewed by a number of experts, including authors of studies identified during the selection and evaluation process.

Checklist for Evaluating Information on an Indicator

The checklist was developed as a guide to evaluating the reliability, validity, and feasibility of information on indicators. An annotated copy of the checklist is included with the procedure following this narrative; this narrative is intended to define the categories and explain the rationale for their inclusion. Categories included in the checklist were organized as follows:

- basic descriptive material,
- reliability and validity,
- results,
- external validity, and
- feasibility of using indicator.

Readers were instructed to note *basic descriptive material* including the name of the indicator; information about the title, author, and publication source of the citation; and descriptions of the study place and population (including patient and provider characteristics) and of the method and measures used in the study. Categories were then provided to assist readers in assessing the reliability and validity of the measures and the study. If the face validity; reliability; and content, convergent, and construct validity of a measure had been established in other studies or in a primary source, readers were asked to provide references to the relevant studies and to evaluate the source material. Readers were asked to note whether observations concerning validity and reliability (and later, feasibility) were made by the authors of the study being evaluated, other reviewers, or the reader.

It has been argued (410,411) that evaluations of quality indicators should focus on measurement issues[2] rather than causal relationships. However, because many of the studies attempting to establish the validity of indicators of quality posit causal relationships, OTA included categories relevant to both types of studies.

Reliability was defined, as it usually is, as the consistency in results of a measure, including the tendency of a test or measurement to produce the same results twice when it measures some entity or attribute believed not to have changed in the interval between measurements. Readers were asked to address the reliability of each measure in the study, with particular attention to the data bases used, because standard data bases are used in many quality studies.

Face validity was defined as being equivalent to intelligibility; that is, the reader was asked to judge (or record, if others had previously evaluated face validity) whether the measure and hypothesized relationships would make sense to the average consumer and provider.

Several of the types of validity included in the checklist—content, convergent, and construct validity—overlap somewhat. As noted by Cronbach, the end goal of validation is explanation and understanding; therefore, the measurement profession is coming around to the view that "all validation is construct validation," and that other types of validation do no more than spotlight aspects of the inquiry (156).

Construct validity is the extent to which a measure measures what it is supposed to measure. McAuliffe, who has written specifically about the validity of indicators of the quality of medical care, points out that the principle underlying content, convergent, and construct validity is to examine, with empirical findings, the consistency of a network of assumptions about the validity of a measure (410). In the broadest sense, then, OTA's entire assessment of indicators can be thought of as validation of indicators of the construct "quality."

Readers were also asked to consider threats to construct validity as traditionally defined. These included inadequate preoperational explication of the target construct; having only one exemplar of the target construct (this would apply to the indicator level); and having dimensions that are irrelevant to the target construct (147). Readers were also requested to note other threats to construct validity.

---

[2] Measurement is "the process by which things are differentiated" (303). Principles of measurement theory have been applied primarily to educational and psychological tests as well as to evaluations of performance (618). Principles of measurement are discussed in the sections on content and convergent validity in this appendix and in the checklist developed by OTA, and explicated further in McAuliffe (410), Nunnally (467), Thorndike (618) and others.

Content validity concerns how representative the sample of items is of the universe it was intended to represent. Content validity depends more on qualitative judgment and does not, by itself, yield a quantitative estimate of the degree of validity (410). To determine content validity, readers were asked to consider: 1) whether the substantive domain of the measure had been adequately specified (e.g., is the measure based on medical knowledge gained through research, clinical experience, and analysis?); and 2) whether adequate scoring rules and procedures for collecting, processing, and analyzing the measure had been developed. Readers were also asked to note how the measure could be improved, according to the author of the study being evaluated, critics, or the reader.

Convergent validity depends on the correlations among two or more measures of a concept, and is another way to help establish construct validity (618). The converse of convergent validity is *discriminant validity*. Discriminant validity would be indicated by much lower correlations between measures of the construct being validated and ones designed to measure some other construct (618). In a systematic approach, a matrix of correlations among measures can be examined. If measures agree with those with which they have been predicted to agree, and disagree with those with which they have been predicted to disagree, the proposed theoretical interpretation (i.e., that those agreeing measure quality) is supported. This multimethod principle must be satisfied by any scientific construct (707).

Convergent validity does not, however, presuppose that one measure is a standard against which other measures should be evaluated. The latter type of validity is concurrent validity. A concurrent study is logical, for example, when an alternative is proposed as a substitute for a measure that is more expensive or difficult to use (156). If construct validity has been established for the more difficult or expensive measure, it may be used as a *criterion* or "gold standard" against which other measures (tests, indicators) are evaluated (207,410). Quality assessment and, as a consequence, OTA's assessment, are hampered by the lack of a criterion for quality against which to validate indicators (410); thus, the checklist was not designed to measure concurrent validity.

Internal validity refers to the extent to which the design of a study contributes to the confidence that can be placed in the study's results. Internal validity is relevant to both measurement studies and studies of causal relationships; it is the extent to which the relationships detected in a study are not spurious (i.e., due to factors not accounted for in the study). Studies of quality indicators rarely use randomized clinical trials and sometimes use voluntary provider-participants; thus, they are frequently open to bias. A number of other threats to internal validity have been enumerated (147,554). The most relevant of these were included in the checklist. Readers were also asked to note when studies did unusual things to improve internal validity.

Statistical conclusion validity is the extent to which research is sufficiently precise or powerful to enable observers to detect effects. Conclusion errors are of two types: Type I is to conclude there are effects (or relationships) when in fact there are not; Type II is to conclude there are no effects (or relationships) when in fact they exist. Readers were asked to describe the analytic method used in the study and to consider the following threats to conclusion validity: 1) whether the sample size was adequate; 2) whether the measures were independent of each other; 3) whether optimal or appropriate statistics were used; and 4) whether controls for case complexity/patient severity were adequate.

External validity is the extent to which the results of a study can be generalized. In evaluating external validity, readers were asked to note factors that would seem to make the results of the study not generalizable across populations, settings, providers, procedures, diagnoses, etc. Inferences concerning external validity in each study were to be compared across studies after the body of literature on an indicator was reviewed.

A section on *feasibility* asked the reader to address whether it was practical to develop information on the quality indicator so that the indicator would be useful for consumers. Readers were asked to consider the intelligibility/understandability of the indicator; the availability of data; the resource consumption involved in data retrieval, analysis, and distribution; confidentiality issues related to the release of information; the corruptibility of data by providers; and the stability of the indicator from year to year. Readers were cautioned that it would be unnecessarily duplicative to fill in the details of the feasibility section for every study; the section was available in every checklist to make it easier to note unusual factors related to feasibility.

For some indicators, readers described the *results* of each study in a technical working paper (see app. A). Included in the description were the unit of analysis used in the study; descriptive information (e.g., for the volume indicator, the actual volume observed for each provider); the format in which the results were described; the actual results as reported in the study; and, if possible, the effect size.

The effect size is a critical component of a quantitative research synthesis; it reduces the results of each study included in the research synthesis to a common metric, allowing comparisons across studies. Effect sizes of various studies can be aggregated and an overall effect size derived. The goal is to obtain a "pure number, one free of our original measurement unit with which to index what can be alternatively called the degree of departure from the null hypothesis of the alternative hypothesis" (137). The effect size is most commonly operationalized as the difference between a treatment (experimental) group and a control group, adjusted (i.e., divided) by the error term; however, the original use of effect size was the average correlation coefficient in a body of studies, and causation is not necessarily implied (137,291,710). Because of wide variations in the way results were specified and because analyses were often not quantified (e.g., analyses of content validity), effect sizes could not be calculated.

Discussion and Implications for Future Research

Most proponents of techniques for systematic literature reviews have extolled the advantages of *"meta-analysis,"* which is typically taken to mean the *statistical* or *quantitative* analysis of a large collection of results from individual studies for the purpose of integrating the findings (254). Meta-analysis so defined involves the development of coding categories to accommodate most of the variation in the literature identified, including both substantive and methodological characteristics (710). These coding categories would be fleshed out quantitatively, so that relationships among variables (measures, constructs) could be explored statistically (584). In part because of the nature of the quality literature, and in part because of resource limitations, OTA was unable to develop such a quantitative scheme. It would be very valuable if future research on quality indicators were to develop and execute a quantitative analysis. Such analyses have considerably enhanced the quality of the debate in other fields (553).

As a necessary precursor to a quantitative scheme, OTA's procedure and checklist might be refined. Given resource limitations, OTA's generic checklist proved to be somewhat cumbersome. The checklist was not easy to use systematically with each type of information available on each indicator. Revising the checklist to make it more relevant to each specific type of indicator would have been useful. In addition, OTA's procedure and checklist were oriented to evaluating and synthesizing empirical studies, and they might be improved to apply more clearly to other types of information encountered when evaluating potential quality indicators (e.g., legal analyses of malpractice awards, administrative rulings on disciplinary actions, professional standards for accreditation, and board certification). This would involve closer attention to criteria for content validity.

In conclusion, OTA found its procedure and checklist for evaluating quality indicators, even with their limitations, extremely valuable. Developing the procedure heightened the awareness of readers to potential biases in the selection of information and the importance of a systematic approach to review. The checklist's explication of requirements for reliability, validity, and feasibility served as a useful guide. The fact that this guide was used fairly systematically across the indicators enhances considerably the confidence that can be placed in OTA's analysis and conclusions.

OTA's Procedure for Evaluating an Indicator of Quality

I. Describe the indicator.
  A. Identify indicator.
  B. State hypotheses about relationship between the indicator and the relevant dimensions of quality of care.

II. Select information to evaluate.
  A. Define the universe of information related to the indicator. (This may be an iterative process.)
  B. Use a combination of techniques to identify citations.
    1. Examine existing reviews.
    2. Search appropriate data bases.
    3. Query experts, especially about unpublished studies.
    4. Add appropriate references cited in the studies obtained.
  C. Acquire citations.
  D. Develop criteria for inclusion and exclusion of citations.
    1. Discard citations that are inappropriate to the topic. Give priority to citations that

test hypotheses about the validity of the indicator.

2. Develop in consultation with OTA and apply any other criteria used for inclusion or exclusion of studies, such as random sampling of all citations obtained.
3. Record citations included and excluded.

III. Evaluate citations selected.
  A. Use the attached OTA checklist to evaluate the citations using one of the following methods:
    1. Use the OTA checklist to evaluate each study.
    2. If it is necessary to reduce the citations evaluated to a more manageable number, take a random sample or develop in consultation with OTA a basis other than random sampling to select studies for application of the checklist.
    3. Before applying the checklist, review all studies to look for patterns in the results and then attempt to explain the patterns. Apply the checklist to all the studies whose results are inconsistent with the hypothesized relationship and dominant results, but to only a sample of the studies with consistent results. Assess whether flaws in methods or differences in approaches, variables, settings, or other factors can explain the inconsistent findings. If no plausible explanations are found for the inconsistencies, apply the checklist to a larger sample of the studies with consistent results.
  B. Apply the checklist to the citations selected.
    1. Identify reviewers.
    2. Train reviewers in the use of the checklist.
    3. Assign two reviewers to rate a sample of the citations.
    4. Evaluate, quantitatively if possible, the reliability of the reviewers' conclusions.
      a. Compute the reliability coefficient at the start of the review process.
      b. Retrain reviewers if reliability problem is identified.
  C. Add categories to the checklist as appropriate for each indicator. For consistency, consult with other reviewers and, if necessary, with OTA before adding categories.
  D. Keep good notes, so that the procedure and checklist can be modified as needed.

IV. Present method and findings in written form.
  A. Present background.
    1. Define the indicator.
    2. State the hypothesized relationship between the indicator and the relevant dimensions of quality of care.
  B. Evaluate the reliability, validity, and feasibility of the indicator as a measure of the quality of care.
    1. Present the findings of the evaluation of the indicator regarding reliability, face validity, content validity, construct validity, convergent validity, internal validity, statistical conclusion validity, and external validity.
    2. Evaluate the feasibility of the indicator as a measure of quality. Consider the use of the indicator by individuals and by organizations in evaluating feasibility.
  C. Analyze the policy implications of the findings and conclusions. Consider the appropriate use of the indicator and any additional research or analysis needed.
  D. If appropriate, present the review methods and results of the studies reviewed in a technical working paper.
    1. State criteria and method used to select citations for inclusion in the analysis. Indicate the number of citations included and excluded.
    2. Describe the review process, including the use of reviewers and evaluation of the reliability of their conclusions.
    3. Describe how the different studies operationalized and attempted to validate the indicator as a measure of quality. Include observations relevant to reliability, validity, and feasibility.
    4. Present the qualitative and quantitative results of the studies. If relationships were found between measures, state the direction and magnitude of the relationships.

CHECKLIST FOR EVALUATING INFORMATION ON AN INDICATOR OF QUALITY

Annotation	Checklist Item
	BASIC DESCRIPTIVE MATERIAL
	Publication:
Presentation is in column format to make the information easily "scannable" across studies/checklists	Title
	Author(s)
	Institutional affiliation(s) of authors
Research findings may vary by date of study	Publication date
Research findings may vary by publication source	Publication source (i.e., name of journal, book, dissertation, other unpublished; provide complete publication information)
	Indicator of Quality Evaluated:
	Did source of information explicitly say it was an analysis of a quality indicator or was the source of a different type?
	NOTE: If the data you are about to review is a subset of the entire publication, it may be helpful to make a note here that there were other purposes for the study. Also state whether you will be reviewing other subparts of the publication.
	Study Population:
Basic description of the study population and place(s) where the study took place, etc. may be necessary to understand causal relationships, differences among studies and issues related to generalizability of study findings	Place where information was gathered

Annotation	Checklist Item
	Study period (time)
	Provider type(s)
	Provider characteristics
	Data source (e.g., database)
	Care characteristics:
	Setting(s) of care
	Procedure(s)
Patient characteristics are important to record because studies may find care/outcome differ by type of patient; or, if all or most studies were only done with one type of patient, results may not generalize to other patient groups	Patient characteristics:
	Age (mean and/or distribution and/or general description)
	Sex
	Ethnic/racial characteristics
	Socioeconomic status
Payment source can be a surrogate for socioeconomic status or age.	Payment source
	Diagnosis(es) (Note: Include criteria for diagnosis(es) in sample selection section under "Internal Validity")
The number of cases in the sample is essential to interpretation of statistical and practical significance	Number of providers/subjects/patients/cases/procedures (all that apply)
	Description of Method and Measures Used in the Study:
	Study design
	Hypothesized relationship(s) among independent and dependent variables and direction of relationships OR Focus of study (if a measurement study).

Annotation	Checklist Item
	Measures:
	Independent variable(s) OR Measure being validated If "causal" study, list and describe all independent measures. (If they have been described fully elsewhere (e.g., your review of another study, a primary source) provide a reference so that the description can be located easily.)
	Primary independent variable OR Measure being validated
	Other independent variables
	Dependent variables OR Comparison ("criterion") measure(s)
	RELIABILITY AND VALIDITY
	Note:
	If the face validity, reliability content, convergent and construct validity of measure have been established in other studies or in a primary source, provide references to the relevant study(ies) and evaluate source material.
	Be sure to note whether issues raised about validity and reliability (and later feasibility) were made by the author(s) of the study, others (e.g., in critiques), or you the reviewer.
Face validity is taken here to be equivalent to intelligibility--that is, would the measure(s) and hypothesized relationships make sense to the average consumer and provider.	Face Validity of Each Measure and of the Hypothesized Relationship Among Variables:
	See above note about avoiding unnecessary duplication.

Annotation	Checklist Item
	Face validity of the independent variable(s) OR Measure being validated Face validity of the dependent variable OR Comparison ("criterion") measure(s) Face validity of hypothesized relationship(s) among variables
Reliability is defined as the consistency in results of a test, including the tendency of a test or measurement to produce the same results twice when it measures some entity or attribute believed not to have changed in the interval between measurements.	**Reliability of Measures and Data Sources:** State whether reliability is addressed in the study. Address the pluses and minuses of the study in terms of reliability for each independent variable (measure being validated) and dependent variable comparison measure). Pay particular attention to the data bases used (e.g., varying completeness of medical records used in study; adequacy of judges used to rate conditions. Reliability of independent variable(s) or measure(s) being validated Reliability of dependent variables(s) (or comparison measure(s)) Address raw data
The principle underlying the following three validation methods is to examine, with empirical findings, the consistency of a network of assumptions about the validity of a measure.	Address calculation of rates, if applicable

Annotation	Checklist Item
This section of the checklist is provided as a guide to evaluating the content validity of measures (indicators), even if the measures and indicators are used in studies professing to evaluate causal relationships. Note that content validity depends more on qualitative judgment and does not, by itself, yield a quantitative estimate of the degree of validity (McAuliffe, 1983).	Content Validity: Note: Apply to "measurement validation studies" or to measure other types of studies. For each measure: 1. Has the substantive domain of the measure been adequately specified? (For example, is the measure based on medical knowledge gained through research, clinical experience, and analysis? If so, describe how. If not, describe basis of measure.) 2. Have scoring rules and procedures for collecting, processing, and analyzing the measure been developed? Are they adequate? How could the measure be improved (according to authors, critics, or you, the reviewer)? SUMMARIZE YOUR VIEW (PRELIMINARY, IF NECESSARY) ABOUT THE CONTENT VALIDITY OF THE MEASURE(S)
Convergent validity depends upon the correlations among two or more measures of a concept. Unlike concurrent validity (which presupposes the existence of a validated criterion), convergent validity does not imply that one measure is a standard against which other measures should be evaluated.	Convergent Validity: (Note: Apply at indicator level or specify whether convergent validity has been/is being/should be evaluated for this measure.)
Construct validity is the extent to which an indicator (measure) performs in theoretically expected ways.	Construct Validity: Consider: 1. whether construct validity is addressed in the study, and 2. the pluses and minuses of the study in terms of construct validity for each measure.
Inadequate operationalization of constructs can result from inadequate preoperational	The following should be considered: Are the constructs operationalized adequately?

Annotation	Checklist Item
explication cf constructs; having only one exemplar of a construct ("mono-operation bias"); or having the operation measure contain dimensions that are irrelevant to the target constructs ("surplus construct irrelevancies") (see Cook & Campbell, 1981, for a fuller discussion)	How may exemplars of the construct are there?
	Are all the dimensions of the measure relevant to the target construct?
	If possible, make a preliminary judgement about the construct validity of the measures. Fuller judgements will probably depend on comparing how measures were operationalized in a variety of studies.
Apart from the reliability and validity of the measures used in a study, the design of a study contributes to the confidence that can be placed in the study's results. Internal validity is the extent to which the detected relationships are not spurious (i.e., due to factors not accounted for in the study).	Internal Validity:
Studies on quality rarely use randomized clinical trials and often use voluntary participatory participants; thus, they are frequently open to bias introduced by the nature of the samples studied.	Consider such factors such as: Sample selection (e.g., consider whether participation was voluntary; consider the criteria for inclusion/exclusion of patients/providers)
Subject loss during the study as a threat to validity has also been called "mortality and	Subject retention during study (i.e., patient, provider)

Annotation	Checklist Item
attrition." In designs in which comparisons are made across subjects, subjects' dropping out of the research is a potential source of bias.	
History refers to the occurrence of historical events that potentially affect the outcome variable of interest. History is a potential source of bias whenever comparisons are made within subjects and whenever the order of observation of research participants is not determined randomly.	History
When observations and ratings of the IVs and DVs (e.g., process and outcome) are made by the same person, that individual's hypotheses, expectancies, or self-interest may affect the ratings. In experimental research, this is known as the experimenter expectancy effect, and is avoided, when possible, by having researchers who are unaware of the research hypotheses, or by other stringent means.	Nonindependence of observations

Annotation	Checklist Item
The fact of being measured can influence subjects' responses. In research designs that involve within subject comparisons and a nonrandom order of treatment exposure, such testing effects are a potential source of bias in estimating effects. The use of archival data avoids such problems if the subjects were not aware of being studied prior to the time data collection began. In some field studies, of course, responses to being studied are desirable (e.g., efforts may be made to reduce infection rates). However, these changes then become a confounding effect in interpreting subsequent data.	"Testing"
Maturation occurs when an observed effect may be due to the respondent's growing older, wiser, stronger, more experienced and the like between measurements and when this maturation is not the treatment of research interest. Maturation is a potential source of bias whenever comparisons are made within subject and the order in which subjects are observed is nonrandom. When subject selection is nonrandom, and maturation differs among "subjects" in the sample, selection bias can interact with maturation bias.	Maturation

Annotation	Checklist Item
Changes in the data collection instrument over the course of the study.	Instrumentation
	Other serious methodological flaws that threaten the internal validity of the study
	Are there unusual things the researcher(s) did to improve the internal validity of the study?
Statistical conclusion validity (sometimes called conclusion validity) is defined as the extent to which the research is sufficiently precise or powerful enough to enable observers to detect effects. Conclusion errors are of two types: Type I is to conclude there are effects (or relationships) when in fact there are not; Type II is to conclude there are no effects (or relationships) when in fact they exist.	**Statistical Conclusion Validity:**
	Analytic method
	Conclusion validity:
	Are measures independent of one another?
Conclusions about the presence or absence of effects (or relationships) compare variation in the dependent (comparisons) variable with other sources of variation in the study.	Are controls for case complexity/patients severity adequate?
	Are optimal or appropriate statistics used?
If a finding is not statistically significant, it may be that the sample size is not large enough for a	Is sample size adequate?

Annotation	Checklist Item
meaningful difference to be detected. The power of the statistical test used can be examined after-the-fact.	
	RESULTS:
	Unit of Analysis (Is unit of analysis appropriate?)
	Descriptive Information Provided in the Results Section
	Format (metric) in which results are described
	Actual Results as Reported in the Study (including levels of significance) described to indicate the direction and magnitude of any relationships
Reduction of individual study results to a common metric allows comparisons across studies.	Effect Size: To be calculated if possible. Analytic method, rationale, and calculations would be shown.
This section would be a preliminary summary of how well done the study is overall. What were the results? Are there alternative explanations for any of them? How serious are the flaws in this study? If more information is needed to make these judgments, it might be good to make a note to get that information.	SUMMARY--RELIABILITY, VALIDITY, AND RESULTS:

Annotation	Checklist Item
	EXTERNAL VALIDITY:
	Factors that would seem to make the results of the study not generalizable across populations, settings, providers, procedures, diagnoses, etc. would be described. Inferences concerning external validity in each study would be compared across studies after the body of literature has been reviewed.
	FEASIBILITY OF USING INDICATOR:
This section addresses whether it is practical to develop information on the quality indicator for consumers.	Note: As with the reliability and validity of measures, it would be unnecessarily duplicative to fill in the details of this section for every study. However, having the section available in every checklist would make it possible to note unusual items (e.g., of possibilities for gamesmanship)
Some indicators/measures (e.g., mortality, volume) will be more understandable to consumers than others (e.g., quality "indexes")	Intelligibility/Understandability (from Face Validity section above)
Judge how readily available the data used in the study under review is to consumers or to those who would develop information on the indicator for consumers (e.g., researchers, employee benefit plans, government programs).	Data Availability
From a policy perspective, a balance between costs (in, for example, time and money) and the reliability and validity of measures will probably need to be struck.	Resource Consumption (time and money involved in data retrieval, analysis, and distribution)
Providers or patients may not wish to relinquish certain information. Some information is	Confidentiality

Annotation	Checklist Item
required by some state or Federal laws (e.g., New York State requires the reporting of in-hospital deaths; the Food and Drug Administration requires reporting of deaths as a result of transfusion errors. Studies may not address this issue, but if they do, or if the reviewer has knowledge from some other source, the issue should be addressed.	
Gamesmanship/corruptibility is the extent to which a provider (or assessor) can manipulate data to make themselves "look good" (or, in the case of diagnostic-related group, for example, increase the reimbursement rate they receive.)	Gamesmanship/ Corruptibility
	Stability of Indicator From Year to Year

SUMMARY--FEASIBILITY:

NOTES

Appendix D

Quality Assessment Activities by Selected Organizations

Various organizations are engaged in activities related to assessing the quality of medical care. This appendix describes the efforts of three groups: the American Medical Association (AMA); the Joint Commission on the Accreditation of Healthcare Organizations (JCAHO); and utilization and quality control peer review organizations (PROs). As a professional organization, a nonprofit accrediting body, and governmental contractors, respectively, these organizations illustrate the diversity of interests involved in quality assessment. They also convey the evolutionary nature of quality assessment, since each group is adopting new approaches.

Quality Assessment Activities of the American Medical Association

To strengthen its commitment to high-quality care, the American Medical Association's (AMA) Board of Trustees created a new initiative on Quality of Medical Care and Professional Self-Regulation. The various elements that make up this initiative are outlined in Report QQ, adopted by the House of Delegates in June 1986 (33).

The AMA Physician Masterfile, currently the most comprehensive source of past and current information on physicians, contains data on every physician practicing in the United States (672). It also includes data for U.S. medical school students and graduates of foreign medical schools who are living in the United States. Information on each physician includes the physician's birthplace, age, address, medical school, residency training, specialty, board certification, hospital affiliation, States of licensure, and any State medical board disciplinary actions. Information is not added to the Masterfile unless verified by a primary source (e.g., State licensing agencies for information on a physician's licensure status and the American Board of Medical Specialties for information on board certification status). The AMA Masterfile is routinely used for verifying physician credentials by hospitals; national, State, and county medical associations; Federal and State agencies; and other organizations. Information on physicians is also available to individual consumers who write to request it.

Disciplinary actions taken by State medical boards that affect a physician's medical licensure are reported to the AMA Masterfile by the Federation of State Medical Boards on a monthly basis (672). To prevent a physician who has lost his or her medical license in one State from obtaining a license in a different State, the AMA sends out "licensure action alert letters." When the AMA receives notice of a final disciplinary action taken against a physician who has held or currently holds multiple State licenses, it automatically alerts the other State licensing boards of the sanctioned physician. The AMA's first licensure action alert letter was sent in January 1985 (673). Since then, the AMA has sent State licensing boards an average of 100 to 120 alert letters (regarding 60 to 70 final disciplinary actions) each month. The AMA also sends alert letters in response to requests for information on or verification of the credentials of a physician, if the physician had a final State disciplinary action on his or her record. These letters advise the requestor to contact for details the appropriate State medical board that took the action.

The AMA's initiative on the Quality of Medical Care and Professional Self-Regulation delineates plans to improve and expand the Physician Masterfile by adding hospital disciplinary actions, malpractice claims and settlement data, and sanctions imposed by the U.S. Department of Health and Human Services (33). The AMA hopes to reduce the amount of time it takes to process a physician credential check to 5 days.

A section of the Health Care Quality Improvement Act of 1986 (Public Law 99-660) mandated the formation of a clearinghouse for information on physicians. The AMA and the Federation of State Medical Boards have formed a partnership in hopes of becoming the designated source of this clearinghouse (673). Data in the mandated clearinghouse include hospital and State disciplinary actions and physicians' paid malpractice claims. The 1986 law requires that hospitals report these data to the clearinghouse. Should the AMA Masterfile become the legal physician data bank, the proposed JCAHO standards to require hospitals to report disciplinary actions to the Masterfile and to use the Masterfile when making staff privilege decisions would become a legal requirement.[1]

---

[1] The national data bank did not receive funding for fiscal year 1988, although it is in the President's budget for fiscal year 1989 (669).

In addition to maintaining the Masterfile, the AMA maintains a file containing information on approximately 70,000 deceased physicians (672). Data in the Deceased Physician Report are made available to State licensing boards to prevent individuals from falsifying their records by using the credentials of a deceased physician.

The AMA plans to take the following steps to encourage the regulation of physicians' behavior by their peers (34):

- review the records of AMA members and expel any physician who has engaged in serious misconduct or has been found to be incompetent;
- publish a comprehensive list of peer review guidelines that will encourage active peer review and is intended to help protect physicians who participate in good faith peer review against liability;
- work with the U.S. Department of Justice to clarify the antitrust laws that impede good faith peer review, the hope being to expand the areas of peer review that can be performed without violating antitrust litigation; and
- assist in defending any physician or medical society that is accused of violating antitrust laws if the litigation resulted from good faith efforts at reporting incompetence.

Because of the increasing need to define and measure the quality of medical care, the AMA, through its Council on Medical Service, has defined eight essential attributes of high-quality care and has provided specific guidelines for quality assessment methods (34). The eight attributes of high-quality care are as follows:

1. It produces the optimal possible improvement in the patient's physiologic status, physical function, emotional and intellectual performance and comfort at the earliest time possible consistent with the best interests of the patient.
2. It emphasizes the promotion of health, the prevention of disease or disability, and the early detection and treatment of such conditions.
3. It is provided in a timely manner, without either undue delay in initiation of care, inappropriate curtailment or discontinuity, or unnecessary prolongation of such care.
4. It seeks to achieve the informed cooperation and participation of the patient in the care process and in decisions concerning that process.
5. It is based on accepted principles of medical science and the proficient use of appropriate technological and professional resources.
6. It is provided with sensitivity to the stress and anxiety that illness can generate, and with concern for the patient's overall welfare.
7. It makes efficient use of health care resources needed to achieve the desired treatment goal.
8. It is sufficiently documented in the patient's medical record to enable continuity of care and peer evaluation.

Favorable outcomes, according to the AMA Council on Medical Service, are an inherent characteristic of high-quality care. The AMA will further develop the council's guidelines for quality assessment methods and will encourage their implementation in professionally conducted quality assessment programs (34). It will also explore the feasibility of developing more specific criteria that can be used to measure the eight attributes of high-quality care.

A patient information brochure on the methods the medical profession currently uses to ensure quality of care and on how patients themselves can evaluate the quality of care they are receiving has been prepared by the Council on Medical Service (37).

The AMA intends to expand its activities relating to geographic variations in the utilization of health care services (266). The AMA publication *Confronting Regional Variations: The Maine Approach* describes an active approach to confronting a situation with many quality implications (39). By supplying feedback to physicians, based on health service utilization data for a specific area, providers can reassess clinical practice patterns, and perhaps improve the quality and efficiency of their services by adjusting inappropriate patterns. Such demonstration projects have also been proposed for Texas, Wisconsin, and Massachusetts (471). Funding for these studies is currently being discussed.

The AMA initiative also calls for the appointment of a commission that is to review the standards for evaluating the clinical performance of medical students and graduates of foreign medical schools (471). The commission is also expected to investigate how medical education could be modified to influence the behavior of physicians.

Quality Assessment Activities of the Joint Commission on the Accreditation of Healthcare Organizations

Since 1951, Joint Commission on the Accreditation of Healthcare Organizations (JCAHO), formerly the Joint Commission on Accreditation of Hospitals, has operated a voluntary accreditation process designed to ensure the quality of medical care services provided in health care organizations. By using structure and process standards that could be evaluated in a survey, the Joint Commission intended to show that JCAHO-accredited organizations have the mechanisms in place to provide high-quality patient care. In 1987, JCAHO accredited approximately 5,000 hospitals and 2,600 other health care organizations, including psychiatric, alcoholism, drug dependence, and mental retardation/developmental disabilities organizations, ambulatory health care organizations, long-term care organizations, and hospices. JCAHO accreditation surveys for home care organizations and managed care organizations are going to be introduced in 1988 (524).

The Current JCAHO Accreditation Process

To be eligible for a JCAHO accreditation survey, a hospital or other health care organization must first meet certain criteria.[2] Among the criteria are having a governing body, an organized medical staff, and a nursing service; providing certain specified services, such as diagnostic radiology services and medical record services; and providing at least one acute care clinical service, such as obstetrics-gynecology or adult psychiatry. These prerequisites prevent health care organizations operating below a minimum level from receiving JCAHO accreditation. Thus, the fact that a hospital has JCAHO accreditation at all, independent of its degree of compliance with specified standards, may itself be an indicator of quality.

The current onsite JCAHO survey process typically lasts from 2 to 15 days, depending on the type and size of the organization. For each JCAHO standard, JCAHO surveyors assign a score on a scale between 1 (best) to 5 (worst), based on the facility's degree of compliance with the provision of the standard. For any score worse than 2, JCAHO surveyors document their reasoning. For hospitals, the individual scores for each JCAHO standard are aggregated into the 8 main categories and 43 elements in shown in table D-1. The JCAHO system for rating the 43 elements is shown in table D-2. For any element that receives a rating below 2, the hospital receives a "contingency."

Depending on the criticality and pattern of elements receiving a contingency, JCAHO may decide to require a written progress report from the organization within a specified period ranging from 1 to 9 months (depending on the issue), may conduct a more focused survey of the facility, or, if the element is particularly crucial, may refuse JCAHO accreditation. In most cases, an institution with a certain number of contingencies will be awarded JCAHO accreditation, with the requirement that the institution correct the deficiencies within a specified time. Each year, 93 percent of the hospitals that JCAHO surveys receive at least one contingency (238). The denial of JCAHO accreditation can result from the overall level of failure of a facility to be in substantial or significant compliance with JCAHO standards and/or from certain patterns of failure in especially critical areas. If JCAHO determines

Table D-1.—Main Categories and Elements of JCAHO Hospital Accreditation Surveys

1. **Laboratory**
    - a. Proficiency testing
    - b. Quality control
    - c. Administrative procedures
    - d. Safety
    - e. Professional staff
2. **Medical records**
    - a. Delinquency
3. **Medical staff**
    - a. Appointment/reappointment
    - b. Clinical privileges
    - c. Direction and staffing
    - d. Organization
4. **Monitoring and evaluation**
    - a. Ambulatory care services
    - b. Anesthesia services
    - c. Dietetic services
    - d. Emergency services
    - e. Home care services
    - f. Nuclear medicine
    - g. Pathological and medical laboratory services
    - h. Pharmaceutical services
    - i. Radiology services
    - j. Rehabilitation services
    - k. Respiratory care
    - l. Social work service
    - m. Special care units
5. **Monitoring functions**
    - a. Medical staff/departmental monitoring and evaluation
    - b. Drug review
    - c. Blood review
    - d. Medical record review
    - e. Pharmacy and therapeutics review
    - f. Surgical case review
    - g. Utilization review
    - h. Infection control
6. **Nursing services**
    - a. Nursing process
    - b. Licensure
    - c. Direction and staffing
    - d. Monitoring and evaluation
7. **Plant, technology, and safety management**
    - a. Life safety
    - b. Safety operations
    - c. Equipment management
    - d. Management of utilities
8. **Quality assurance programs**
    - a. Governing body/management support
    - b. Written plan
    - c. Quality assurance results a determinant of clinical competence/privilege
    - d. Evidence of actions

SOURCE: Joint Commission on Accreditation of Healthcare Organizations, "Hospital Accreditation Program: Accreditation Decision/Contingency Criteria," Chicago, IL, 1987.

---

[2] These eligibility criteria may differ for different types of health care organizations.

Table D-2.—System Used To Rate Elements and Assign Contingencies in JCAHO Accreditation Surveys

Extent of institution's overall compliance with standards in an element	Rating	JCAHO's contingency response[a]
Substantial compliance	1	Accreditation/no contingency
Significant compliance	2	Accreditation/no contingency
Partial compliance	3	Accreditation with contingency[b]
Minimal compliance	4	Accreditation with contingency[c]
No compliance	5	Accreditation with contingency[c]
Not applicable	NA	Not applicable

[a] The contingency responses listed below are accompanied by JCAHO's recommendations for improvements that must be made within a specified time to bring the institution into full compliance with JCAHO requirements.

[b] The institution must submit a written progress report to JCAHO in a specified time period. The contingency score for the element may not be aggregated with other contingency scores to warrant a focused survey of the institution, but if a focused survey is conducted, the element must be included.

[c] The contingency score for the element may be aggregated with other contingency scores to warrant a focused survey of the institution or may result in nonaccreditation by JCAHO.

SOURCE: Joint Commission on Accreditation of Healthcare Organizations, *AMH/87: Accreditation Manual for Hospitals* (Chicago, IL: 1987).

that an organization may be denied accreditation, the facility is specially reviewed and given more individualized attention in an effort to bring it into compliance with the standards. Only 1 to 2 percent of JCAHO-surveyed hospitals each year do not come into substantial compliance in a timely fashion and are denied JCAHO accreditation (238).

Implementing New or Revised JCAHO Standards

Revisions in JCAHO standards are developed by JCAHO with the assistance of consultants or special task forces, and then forwarded to professional and technical advisory committees. If these advisory committees recommend the revisions, the proposed changes are sent to the Standards and Survey Procedures Committee of JCAHO's Board of Commissioners along with a request that the Standards and Survey Procedures Committee approve the revisions and allow them to be reviewed further by 2,000 to 5,000 professional organizations, individuals, and other interested parties, including a percentage of the accredited organizations.

After the reviewers comments are analyzed, JCAHO's Department of Standards and the consultants or special task force may revise the standards. The proposed standards are presented again to the professional and technical advisory committees and to the Standards and Survey Procedures Committee. Additional field reviews are undertaken, depending on the extent and nature of the revisions to the proposed standards. After all revisions have been made, the final proposed standards are submitted to JCAHO's Board of Commissioners to adopt for use in JCAHO accreditation surveys (559).

Elements of new or revised JCAHO standards are occasionally placed in "implementation monitoring." Affected institutions are given additional time for effectively implementing a new or revised standard while JCAHO surveys and monitors their progress toward compliance, but the institution's level of compliance with the standard does not affect JCAHO's accreditation decision. No less than annually, any standards in implementation monitoring are reviewed, and if institutions have had sufficient time to successfully implement the new or revised standards, the standards will be taken out of implementation monitoring and the organization's compliance will be considered in JCAHO's accreditation decision (559).

JCAHO's 1986 "Agenda for Change"

In September 1986, JCAHO announced an "Agenda for Change" that signified a major redirection in its approach to quality assessment (523). The principal initiative of this agenda centers around a new approach to the current JCAHO survey and accreditation process. In the past, JCAHO has relied exclusively on structure and process standards to evaluate the capability of an organization to provide high-quality care.[3] Project Objective I of JCAHO's Agenda for Change calls for the development of indicators to assess the actual clinical performance of the organization, including the outcomes of the medical care it provides.

JCAHO believes that with recent advances in health care research methods, it is now possible to monitor an organization's clinical performance and outcomes more precisely, moving beyond answering the basic question, "*Can* this organization provide quality health care?" to answer the question, "*Does* this organization provide quality care?" (329)

---

[3] In the early 1970s, responding to criticism that it placed too much emphasis on physical and administrative structures, the Joint Commission on the Accreditation of Hospitals began to require outcome-oriented hospital quality review programs (333). By 1976, the Joint Commission had developed an outcome-oriented method to audit medical care that was based on retrospective review using preestablished criteria. This method (the Performance Evaluation Procedure for Auditing and Improving Patient Care) could be applied to any diagnosis or surgical procedure. In 1979, the Joint Commission eliminated the medical audit requirements because while being costly, they often focused more on the data collection process than on problem solving (10). Furthermore, the medical audit requirements focused on already suspected problems, rather than on identifying problems and opportunities to improve care. The requirements were replaced with an organization-wide quality assurance system.

With assistance from expert groups, task forces, medical specialty societies, and accredited institutions, JCAHO plans to develop valid indicators of clinical performance of health care organizations. The indicators will be selected from clinical areas associated with high-volume/high-risk and/or potentially problematic care (523). Task forces have already proposed clinical indicators for hospital-wide care and for obstetrics and anesthesiology, and in 1987, pilot tests of the indicators began in 17 hospitals (11). Some of the indicators will be aggregated rates and others will be single sentinel events; they will be used to evaluate both diagnostic and treatment activities. Structure, process, and outcome indicators will be selected so as to be applicable to organization-wide reviews, cross-departmental reviews, and specialty-specific reviews. Examples of organization-wide clinical indicators include mortality rates of patients with specified medical conditions; examples of cross-departmental indicators for surgical departments include specific complications for specified surgical procedures. JCAHO asserts that the clinical indicators of quality developed will not measure the quality of care directly, but rather will serve as "flags" to identify care that requires further analysis and review (329). By identifying potential quality-of-care problems and areas in which care can be improved, JCAHO and the health care institutions can focus directly on those areas of patient care that are in most need of attention.

Another aspect of JCAHO's "Agenda for Change" are revisions in the organizational assessment of health care institutions. Project Objective II includes the development of valid intra-organizational indicators, using organizational research findings and the advice of experts. These indicators could be used to improve the monitoring of the organizational functions such as planning, resource allocation, leadership, and evaluation that are believed to influence the quality of care most directly (329).

The comparison of different organizations using clinical indicators could be improved by a valid method to adjust for differences in the severity-of-illness of the patients that the organization serves. Project Objective III of JCAHO's "Agenda for Change" calls for the development of a method to adjust for patient differences so that equitable comparisons can be made among institutions. JCAHO, along with the help of experts in this area, plans to examine current severity-adjustment methods, and if necessary, to modify or create new methods that more adequately account for the confounding effects of patient variables on measures of institutional performance. With the use of a valid severity-adjustment method, an institution could compare its own results for an indicator to the results of other institutions or to a standard norm, without confusion caused by differences in the severity of illness among the patient populations (329).

Project Objective IV of JCAHO's "Agenda for Change" concerns the assessment of current institutional data bases and monitoring systems to test their applicability to the collection and analysis of data for clinical and organizational indicators of an organization's performance. JCAHO will provide technical assistance to those institutions that must develop a clinical and organizational data collection process that is more outcome oriented. JCAHO will also continue to provide assistance with the establishment and modification of appropriate internal quality assurance systems. The extent to which JCAHO data reporting requirements are coordinated or could be tailored to be coordinated with other external data reporting requirements, such as those of the Health Care Financing Administration (HCFA) Medicare data set, will also be determined (523).

The creation of an ongoing interactive monitoring system between the JCAHO and the accredited institutions is another aspect of Project Objective IV. Rather than only conducting onsite surveys of each health care organization every 3 years, JCAHO hopes eventually to collect data on the indicators from each organization three to four times per year (119). At these regular intervals, JCAHO will collect the data relative to the specified indicators of clinical performance and organizational performance that the organization's departments will be collecting continuously. These data would be submitted to the JCAHO either in writing, by diskette, by data tape, or by modem. After the JCAHO processes the information gathered, it plans to provide feedback, in the form of aggregate and facility-specific evaluations of clinical and organizational performance, including outcomes, to each health care facility. With these new data, each institution could then compare its performance to the standing of other similar facilities or to external expectations (based on national and regional performance standards). Continual feedback from JCAHO could complement an institution's own self-monitoring process and serve as an "early warning system" to draw attention to an area needing prompt evaluation. JCAHO plans to analyze further the issues of cost and feasibility of this ongoing interactive monitoring (329).

To accommodate the intensive monitoring system and the new focus on clinical and organizational indicator data, the JCAHO plans to revise the accreditation survey and the accreditation decisionmaking process. Project Objective V of the JCAHO's "Agenda for

Change" addresses the assurance of the validity, reliability, and utility of the new data to be accumulated by each health care organization. Surveyors will evaluate the organization's analysis of problem areas and assess the effectiveness of actions taken to resolve recognized problems. JCAHO will also examine how information from surveys and from the ongoing monitoring activities will be integrated into the accreditation decisionmaking process (329).

JCAHO realizes that with such an extensive data base on institutional performance and because of increasing demands for public accountability, confidentiality and disclosure policies must be discussed. Although currently JCAHO upholds strict confidentiality policies, it speculates that there is the potential for the release of aggregate data, but there are no current plans to release institution-specific data (523).

JCAHO plans to gradually implement the objectives of the "Agenda for Change" first in pilot tests and then in stages for accredited organizations. During the developmental process, JCAHO plans to monitor closely the capabilities of the health care institutions. During 1988, development of clinical indicators will begin for cardiovascular, trauma, oncology, and surgical care, for long-term care, and for mental health services. Implementation is scheduled to begin in 1989 with hospitals, with full implementation scheduled for the early 1990s first for hospitals, and then subsequently for psychiatric, ambulatory, and hospice services (329).

Quality Assessment Activities of Peer Review Organizations

Utilization and quality control peer review organizations (PROs) are federally mandated under the Tax Equity and Fiscal Responsibility Act of 1982 (Public Law 97-248) to monitor the quality of medical care provided to Medicare beneficiaries.[4] To receive payment under Medicare's hospital payment system based on diagnosis-related groups (DRGs), hospitals are required by the Social Security Act of 1983 (Public Law 98-21) to enter into agreements with PROs. PROs are mandated to review the care these hospitals provide to Medicare patients with the purpose of ensuring that the services are medically necessary, are provided in the most appropriate setting, and meet professionally recognized standards of quality medical care. Under the direction of HCFA of the U.S. Department of Health and Human Services, PROs are able to deny payment for inappropriate services and to take necessary action to correct unacceptable medical practices (535).

HCFA enters into contracts with 54 PROs geographically distributed across the country. The District of Columbia, Puerto Rico, the Virgin Islands, Guam and American Samoa, and each of the 50 States are considered separate PRO areas. To qualify as a PRO, an organization must demonstrate either 1) sponsorship by at least 10 percent of the physicians practicing in the review area, or 2) physician accessibility, i.e., the involvement of at least one physician in every generally recognized specialty in the area (42 CFR 462.102-462.103).[5] Third-party payers can obtain PRO contracts only if it is determined that no eligible organization other than a payer organization is available.[6] In 1985, 41 PROs were supported by a State medical association (158).

The number of personnel working full time in each PRO varies depending on the caseload in the PRO's area. The staff includes mainly nurses, medical-record analysts, clerks, secretaries, and financial managers. Physicians are usually involved on a part-time basis as first-line physician reviewers, consultants, or members of the board of directors. Physician reviewers must have active admitting privileges in one or more hospitals in the area; consultants must be physicians in active practice but do not necessarily have to have admitting privileges (e.g., anesthesiologists, pathologists, and radiologists).

PRO Contracts

Through a competitive bidding process, HCFA, since 1984, has awarded and renegotiated PRO contracts every 2 years.[7] The scope and structure of PRO reviews are delineated in each PRO's contract. Specified criteria in each contract reflect federally mandated objectives for PRO review, provisions specified by HCFA, and particular quality and admissions objectives specific to each of the 54 PRO areas.[8] Each PRO is required by HCFA to propose area-specific objectives, used as measurable targets to be reached during the 2-year contracts. A PRO's performance is evaluated by HCFA regional offices and the HCFA central office on the basis of how successfully the PRO has met its stated objectives (429). HCFA's evaluations are also used for determining PRO contracts for the following cycle. Although PRO contracts are applicable only to the review of Medicare patients, PROs are encouraged to enter into similar contracts with Medicaid and other third-party payers.[9]

---

[4] The PRO program was established as the successor to the Professional Standards Review Organizations program, which had been established by the Social Security Amendments of 1972 (Public Law 92-603). For more information on the Professional Standards Review Organizations program, see K.N. Lohr, "Peer Review Organizations (PROs): Quality Assurance in Medicare" (382).

[5] These organizations must have letters of support (written by other physicians in the area) establishing that they are representative of the specialty, as it is practiced in the PRO area (83).

[6] As of January 1988, the only PRO with a contract held by a third-party payer was the Hawaii PRO.

[7] The Omnibus Budget Reconciliation Act of 1987 (OBRA-87)(Public Law 100-203) mandates that PRO contracts in the next set are to be renegotiated every 3 years.

[8] The PRO contracts for Maryland, New Jersey, the Virgin Islands, Guam, American Samoa, and in the North Mariana Islands have differed with regard to specific objectives, because in these areas, Medicare does not pay for beneficiary inpatient care on the basis of DRGs; these areas have held waivers from the national Medicare program and have been regulated by alternative payment systems (630).

The first round PRO contracts, which became effective between July and November 1984, covered the 2-year period 1984-86. These contracts focused primarily on the detection by PROs of inappropriate utilization and payment patterns. Specifically, PROs were expected to reduce unnecessary hospital admissions, to ensure that Medicare payment rates were based on diagnostic and procedural information contained in patient records, and to ensure that Medicare patients were not readmitted within 7 days of discharge as a result of premature release from the hospital (535). The PRO contracts for 1984-86 were also to include area-specific admission[10] and quality objectives (see table D-3). PRO were allowed to choose the procedures and conditions on which to focus both admission and quality objectives.

The second round of PRO reviews, beginning in July 1986 and covering a 2-year period through 1988,[11] have been more focused on quality-of-care issues (535). Provisions of the Omnibus Budget Reconciliation Act of 1986 (OBRA-86) (Public Law 99-509) required PROs to review health care provided to Medicare beneficiaries enrolled in health maintenance organizations (HMOs) and competitive medical plans (CMPs), but these provisions were not reflected in the 1986-88 PRO contracts; contracts used to implement HMO/CMP review by PROs were implemented in mid-1987 (see discussion below).

Table D-3.—Admissions and Quality Objectives for PROs in the 1984-86 PRO Contracts

Admissions and quality-related objectives in the 1984-86 PRO contracts were as follows:

Admissions objectives

1. To reduce admissions for procedures that could be performed safely and effectively on an ambulatory basis.
2. To reduce inappropriate or unnecessary admissions or reducing invasive procedures for specific DRGs, practitioners, or hospitals.

Quality objectives

1. Reduce unnecessary hospital readmissions resulting from substandard care provided during the prior admission.
2. Assure the provision of medical services which, when not performed, have significant potential for causing serious patient complications.
3. Reduce avoidable deaths.
4. Reduce unnecessary surgery or other invasive procedures.
5. Reduce avoidable postoperative or other complications.

SOURCE: P.E. Dans, J.P. Weiner, and S.E. Otter, "Peer Review Organizations: Promises and Pitfalls," *New England Journal of Medicine* 313(18):1131-1137, 1985.

Table D-4 compares PROs' 1984-86 and 1986-88 scopes of work. As in the first contract period, the medical records reviewed by the PROs in the second contract period are obtained from the fiscal intermediary payment claims for inpatient hospital care. The criteria and specified percentages of cases to be reviewed, however, have been changed in the more recent contracts to reflect the new quality-of-care focus. The PRO scope of work for the 1986-88 contract period includes several new review requirements (659):

- apply generic quality screens to all inpatient cases reviewed in order to identify potential quality problems;[12]
- review hospitals identified because of unexplained statistical outliers in the HCFA data on high mortality rates or utilization patterns;
- review each case selected by the PRO for retrospective review for the appropriateness of discharge;
- develop and implement a community outreach program to educate beneficiaries about PRO review and Medicare rights.

The 1986-88 contracts have included national objectives, which are established by HCFA, and area-specific objectives, which are proposed by each PRO under guidelines specified by HCFA. All objectives are physician or hospital specific. PROs' 1986-88 scope of work stipulates that the following cases are to be reviewed retrospectively:

- a 3-percent random sample of all discharges per hospital;

---

[9] If State Medicaid programs contract with the local PRO for the review of medical care provided to Medicaid beneficiaries, they receive a 75-percent Federal reimbursement, as opposed to a 50-percent reimbursement for contracting with an outside organization (83).

[10] A review of the PRO contract objectives revealed large differences in the proposed reduction targets. Although the PRO contracts for Florida, Georgia, and Iowa each specified a reduction in hospital admissions for lens procedures, Florida targeted its reduction rate at 76 percent, Georgia specified a 25-percent decrease, and Iowa proposed only a 10-percent reduction (474).

[11] The termination dates for the PRO contracts in 1988 differ because of the range in contract initiation dates in 1986.

[12] The following six categories of screens are applied to every case reviewed in order to identify potential quality problems: 1) adequacy of discharge planning, 2) medical stability of patient at discharge, 3) deaths that may indicate poor-quality care, 4) nosocomial infections, 5) unscheduled return to surgery, and 6) trauma suffered in the hospital (see ch. 5 for more details on generic screens).

Table D-4.—Comparison of PROs' 1984-86 Scope of Work and PROs' 1986-88 Scope of Work

Category	1984-86 Scope of work	1986-88 Scope of work
Objectives	Three admissions objectives and five quality objectives, all proposed and validated by PROs; very limited areas for focusing objectives	Five objectives based on PRO data from first 90 days of generic quality screen review.[a] HCFA-identified mortality and utilization outliers. Broader objectives
Random samples	Review a 5-percent sample of all hospital admissions; 3-percent to 100-percent sample of inpatient hospital records for DRG validation (based on number of hospital discharges)	Review a 3-percent random sample of all prospective payment hospital discharges (including, for the first 6 months of PRO contract, all cases with a 1- or 2- day hospital stay)
Preadmission review	Review cases involving any of five procedures proposed by PRO	Review cases involving cardiac pacemaker implants or reimplants plus four procedures proposed by PRO
Cases involving cardiac pacemaker implants or reimplants	Review 100 percent of cases retrospectively	Review 100 percent of cases preadmission (see above)
Transfers	Review all transfers from a prospective payment hospital to another hospital exempt unit or swing bed	Same, but lower percentage of cases are reviewed
Readmissions	Review all readmissions within 7 days of discharge from a PPS hospital	Review all readmissions within 15 days of discharge from a PPS hospital
Medicare code editor	Review 100 percent of nine diagnoses specified by HCFA	Same
Review cases in specific DRGs	Review all cases in DRG 468 (unrelated operating room procedure); DRG 462 (rehabilitation) was added during the contract period	All cases in DRG 468 (unrelated operating room procedure), DRG 462 (rehabilitation), and DRG 088 (chronic obstructive pulmonary disease)
Day and cost outliers[b]	Review 100 percent (reduced to 50 percent during the contract period)	Review a 50-percent sample
Cases involving percutaneous lithotripsy	Not in contracts	Review all claims for percutaneous lithotripsy in hospitals that have an extracorporeal shock wave lithotripter
Validation of objectives	Not in contracts	Review a sample of discharges within a 3-month period to validate PRO's individually negotiated performance objectives
Hospital notices of noncoverage to beneficiaries	Review 100 percent where patient disagrees. 100 percent where patient is liable for charges for services rendered after notification of noncoverage. 10 percent of remaining.	Same
Specialty hospital review	Proposed by each PRO	Review 15 percent of discharges
Admission pattern monitoring	Discontinued during contract	Not in scope of work
Intensified review[c]	Trigger: 2.5 percent of cases reviewed or three cases per hospital (whichever is greater)	If denial associated with 1 department or physician, review increased to 100 percent Trigger: 5 percent of cases reviewed or six cases (whichever is greater). If denial associated with one department or physician, review increased to 50 percent (first quarter) or 100 percent (two or more consecutive quarters)
Community outreach	Not in contracts	All PROs to propose program

[a] Each PRO determines its own specific targets for these objectives according to potential problems of quality of care revealed from the first 90 days of generic quality screen review.

[b] A day outlier is a case in which a hospital seeks payment for days in the hospital exceeding, by a specified amount, the average length of stay paid under Medicare's prospective payment system (PPS). A cost outlier is a case in which a hospital seeks payment for medical care expenses exceeding, by a specified dollar amount, the average level of payment paid for that DRG.

[c] This is a more focused review triggered by a certain percentage of denials for a specific physician or hospital department (often revealed by physician and hospital profiles).

SOURCE: U.S. Department of Health and Human Services, Office of the Inspector General, Office of Analysis and Inspections, "The Utilization and Quality Control Peer Review Organization (PRO) Program," draft report, Control No. OAI-01-88-00570, Washington, DC, February 1988.

- all readmissions within 15 days of discharge from a prospective payment hospital;
- all transfers from one PPS hospital to another. A sample of transfers from a PPS hospital to PPS exempt swing beds, alcohol/drug abuse units, psychiatric units, and rehabilitation units;
- a 50-percent sample of day outliers and cost outliers;
- all cases with DRG assignment for rehabilitation (DRG 462), unrelated operating room procedure (DRG 468), and chronic obstructive pulmonary disease (DRG 088).[13]
- all cases in which the patient disagrees with a notice of non-Medicare coverage by a hospital or in which the patient is liable for the charges for non-Medicare coverage. All cases in which the physician disagrees with hospital notice of non-Medicare coverage. The PRO also reviews 10 percent of all other cases where notices of non-Medicare coverage have been issued.
- a random sample of 15 percent of discharges from PPS-exempt hospitals.
- all cases for percutaneous lithotripsy in hospitals with an extracorporeal shockwave lithotripter and cardiac pacemaker implants or reimplants.
- all cases in which a covered level of care occurs during a hospital admission that the hospital had determined originally to be a noncovered hospital stay.
- all cases that the fiscal intermediary refers to the PRO for a medical necessity determination.

HCFA also requires in the 1986-88 contracts that PROs review cases involving nine specified diagnoses before Medicare payment is provided.[14] In addition, PROs are to perform preadmission review for five procedures, a review for all cases involving the implantation or reimplantation of cardiac pacemakers and four other procedures chosen by each PRO. The four PRO-selected procedures are based on criteria delineated by HCFA and have been designated in the PRO contract. PROs will be required to perform 100-percent preadmission and preprocedure review for 10 different elective surgical procedures, a provision under the Consolidated Omnibus Budget Reconciliation Act (COBRA) of 1985 (Public Law 99-272). Guidelines for this provision have not yet been implemented (83).

Each PRO must also develop specific goals based on the following objectives (83):

1. Eliminate adverse outcomes (including premature discharges) by focusing on providers and/or practitioners and by focusing on DRGs;
2. Reduce unnecessary admissions and/or procedures by provider and/or practitioner and by focusing on DRGs.

Each PRO has determined its own specific targets for these objectives according to potential problems of quality of care revealed from the first 90 days of generic quality screen review, HCFA-identified outliers,[15] or other identified problem areas. Altogether, the cases selected for PRO review have included approximately 25 percent of all hospital discharges for Medicare.

Additional PRO duties have been mandated under COBRA and OBRA-86, but they have not yet been incorporated into the 1986-88 PRO contracts. COBRA allows PROs to deny payment for care of substandard quality as identified through criteria developed under HCFA guidelines. As part of the preadmission review for specific elective surgeries, COBRA also allows PROs to require second opinions if warranted.

PROs' 1988-90 scope of work includes the new requirements mandated in OBRA-86 and in the Omnibus Budget Reconciliation Act of 1987 (OBRA-87) (Public-Law 100-203). Rather than concentrating solely on inpatient hospital care, the 1988-90 scope of work focuses on the continuum of patient care. The third round of PRO contracts will include a requirement that PROs review all hospital readmissions within 31 days of discharge. PROs will also be required to review the intervening care delivered to a percentage of Medicare beneficiaries with hospital readmissions. HCFA's proposed generic quality screens used by PROs for reviewing inpatient hospital records have been revised for the third scope of work (table D-5).[16] The new generic quality screens include a 7-page *Generic Quality Screens Guideline* to clarify the criteria for determining potential quality-of-care problems (see ch. 5).

PROs' 1988-90 scope of work also contains a requirement that PROs review the quality of services among a variety of alternative settings, including ambulatory surgical centers,[17] hospital outpatient departments,[18] and nursing homes. PRO reviews will include all written complaints by Medicare beneficiaries about the quality of services provided in skilled nursing facilities, home health agencies, and hospital outpatient areas.[19] Pilot studies for reviewing the quality of services delivered in physicians' offices is scheduled to begin in January 1989 (83).

---

[13] These DRGs have a high level of payment and tend to be miscoded or abused by hospitals (83).

[14] These diagnoses include diabetes mellitus, noninsulin-dependent and insulin dependent; impacted cerumen; benign hypertension; left bundle branch hemiblock; other left bundle branch block; right bundle branch block; elevated blood pressure without diagnosis of hypertension; and other unspecified complications of medical care not elsewhere classified (651).

[15] HCFA provides each PRO with lists of hospitals in the PROs area that have been identified as having mortality rates that vary significantly from national norms. The PROs were required to evaluate the outliers in their area determined by 1986-87 data. They are not required to perform any focused reviews of outliers revealed in the 1987-88 data.

[16] These new generic quality screens may undergo additional revisions before final implementation.

[17] The review of medical services provided in ambulatory surgical centers and outpatient surgery hospital departments will be incorporated into PRO contracts for those contracts entered into or renewed before January 1, 1987. As of February 1988, Pennsylvania and Massachusetts are the only States that are reviewing these ambulatory settings (83).

[18] Tentatively, PRO reviews of hospital outpatient departments and nursing homes will begin towards the end of 1988 (83).

Table D-5.—HCFA's Proposed Generic Quality Screens for Reviewing Inpatient Hospital Records[a]

- *1. **Adequacy of discharge planning**
  No documentation of discharge planning or appropriate followup care with consideration of physical, emotional, and mental status needs at time of discharge.
- 2. **Medical stability of the patient**
  - a. Blood pressure within 24 hours of discharge (systolic less than 85 or greater than 180; diastolic less than 50 or greater than 110)
  - b. Temperature within 24 hours of discharge greater than 101° F (38.3° C) oral, greater than 102° F (38.9° C) rectal
  - c. Pulse less than 50 (or 45 if the patient is on a beta blocker), or greater than 120 within 24 hours of discharge
  - d. Abnormal diagnostic findings which are not addressed and resolved or where the record does not explain why they are not resolved
  - e. Intravenous fluids or drugs after 12 midnight on day of discharge
  - f. Purulent or bloody drainage of wound or open area within 24 hours prior to discharge
- 3. **Deaths**
  - a. During or following any surgery performed during the current admission
  - b. Following return to intensive care unit, coronary care unit, or other special care unit within 24 hours of being transferred out
  - c. Other expected death
- *4. **Nosocomial (hospital-acquired) infection**
- 5. **Unscheduled return to surgery**
  Within same admission for same condition as previous surgery or to correct operative problem
- 6. **Trauma suffered in the hospital**
  - a. Unplanned surgery which includes, but is not limited to, removal or repair of a normal organ or body part (i.e., surgery not addressed specifically in the operative consent)
  - *b. Fall
  - c. Serious complications of anesthesia
  - d. Any transfusion error or serious transfusion reaction
  - *e. Hospital-acquired decubitus ulcer and/or deterioration of an existing decubitus
  - f. Medication error or adverse drug reaction (1) with serious potential for harm or (2) resulting in measures to correct
  - g. Care or lack of care resulting in serious or potentially serious complications

"Optional Screen"
Medication or treatment changes (including discontinuation) within 24 hours of discharge without adequate observation

[a] The PRO scope of work for 1988-90 includes a scoring system to reflect differences in severity of potential quality problems. For items marked with an asterisk in this table, the PRO reviewer is to record the failure of the screen, but need not refer potential severity Level I quality problems to a physician reviewer until a pattern emerges.

SOURCE: U.S. Department of Health and Human Services, Health Care Financing Administration, Health Standards and Quality Bureau, *1988-1990 PRO Scope of Work* (Baltimore, MD: Apr. 1, 1988).

Reviews of Health Maintenance Organizations and Competitive Medical Plans

Contracts for the review of HMOs and CMPs, mandated in COBRA, were implemented between June and November 1987.[20] All but one HMO/CMP contract have been awarded to existing PROs (428). These contracts require the review of the quality of care delivered to Medicare beneficiaries in HMOs and CMPs. The criteria delineated in the contracts for HMO/CMP review, however, are somewhat different from the objectives contained in the PRO contracts for inpatient hospital review. Each case picked for HMO/CMP review may undergo inpatient review, ambulatory review, and/or post-hospital review (644). The selection of cases for HMO/CMP review is based on the following elements:

- Random sample of 13 conditions, determined by HCFA to be conditions that when leading to inpatient hospital care may be indicative of poor-quality ambulatory care;
- Focused review of ambulatory care services (to begin after the first 6 months of the contract)[21];
- 3- to 6-percent random sample of hospital discharges;
- readmissions within 30 days of discharge from an acute care hospital;
- patient transfers to other hospitals; and
- nontrauma deaths.

An initial analysis of an HMO/CMP's internal quality assurance system determines whether an HMO/CMP will undergo limited or basic review of these elements. When poor review findings exceed certain threshold levels, an HMO/CMP is reassigned to an intensified level of review. Each level of review evaluates cases according to the same criteria. However, the limited review plan evaluates a lower percentage of cases than the basic plan, and the intensified plan analyzes the highest percentage of cases (644).

---

[19] The review of beneficiary complaints was implemented via modification to the 1986-88 PRO contracts (83).

[20] COBRA initially authorized PROs to review the services provided by HMOs and CMPs. This legislation, however, was amended by provisions in OBRA-86. OBRA-86 allowed HCFA to contract for reviews of HMO and CMP services with entities other than PROs on a competitive basis, but these contracts were limited to no more than half of the States, covering no more than half the Medicare HMO/CMP enrollment (627).

[21] The contractor has 6 months from the effective date of the contract to develop and submit a methodology for performing focused review (e.g., by provider, by medical condition) of ambulatory care.

The PRO Review Process

The patient records needed for retrospective review by PROs are identified from Medicare hospital claims submitted to a fiscal intermediary for payment. The fiscal intermediary sends the PRO the data tape for all claims made within a specific time period, usually a month. The PRO analyzes this information with a computer program that flags the specific cases to be reviewed, according to the criteria and specified percentages of cases described above. To obtain the patient records that correspond to the flagged claims, the PRO requests copies of the records (within 30 days) from the hospital, or PRO personnel may go to the hospital to review the records on-site. Physicians are required to notify the PROs of the cases that require preprocedure review (83).

Each record identified for review undergoes five different basic reviews by PRO nurse reviewers. These initial reviews include generic quality screen reviews, admissions reviews, discharge reviews, DRG validation, and items/services coverage reviews. Nurse reviewers use explicit criteria, developed by the PRO, to determine potential quality-related or utilization problems. Should one of these reviews detect a potential problem, the records are referred to a PRO physician adviser for further review (199). Potential quality problems not detected by one of the five reviews, e.g., mismanagement of the case, may be discovered by the initial nurse reviewer based on his or her medical judgment. In this case, the medical record would also be referred to a physician adviser. If the initial reviewer can determine that a case failing one of the generic quality screens is not actually a quality problem, the case is not referred to a physician adviser (627).

A physician reviewer will conduct a more indepth examination of the medical record, based on his or her clinical judgment, to determine whether there actually is a problem. The review process also allows the attending physician and hospital an opportunity to discuss the specifics of the case in question. These discussions often reveal unique characteristics of the case that explain why it may have failed the initial screens. Most cases of potential problems are resolved this way (164).

If the physician reviewer determines after the discussions that the care provided was not medically necessary or that it should have been provided in another setting, a payment denial notice is sent by the PRO to the beneficiary, physician, provider, and fiscal intermediary. If the physician reviewer identifies a quality-of-care problem that is not cleared up after discussing the case with the patient's physician, the PRO will initiate appropriate interventions.[22] These interventions may include physician education through a continuing medical education program, a corrective action plan, intensified review of the physician and hospital, or the initiation of a sanction review (627).

The sanction review process is initiated if other interventions have not corrected the problem or if the quality problem has been determined to be a substantial or a gross and flagrant violation.[23] This sanction process may result in exclusion from the Medicare program or the imposition of monetary penalties (360) (see ch. 6 for a further description of the PRO sanction process).

PROs review the care provided by nearly 7,000 hospitals and 450,000 physicians (164). During the 1986-88 scope of work, PROs took some form of quality intervention, short of initiating the sanction process, against 16,823 physicians and 1,376 hospitals (535). From 1985 through September 1987, 79 sanctions were imposed by the Office of the Inspector General as a result of PRO recommendations: 53 physicians and 1 hospital were excluded from the Medicare program, and civil monetary penalties were imposed on 24 physicians and one hospital (164).

Physician and Hospital Profiles Produced by PROs

PROs also use the data collected from medical record reviews to produce physician and hospital profiles. These profiles include data on denial rates, mortality rates, and review findings on quality and admissions objectives. The PROs analyze these profiles to compare patterns of care by similar providers and current patterns with previous patterns. In addition, the profiles are used to identify patterns of care among physicians and hospitals that deviate from established criteria and standards (627). The identification of an aberrant pattern of care may trigger a PRO's evaluation of a larger sampling of records from the physician or hospital in question.

---

[22] COBRA allows PROs to issues denial notices for substandard quality of care, but HCFA, as noted earlier, has not yet implemented regulations regarding these types of denial notices.

[23] As noted in ch. 6, a substantial violation refers to a pattern of care that is inappropriate, unnecessary, or does not meet recognized professional standards of care, or is not supported by the necessary documentation of care, as required by the PRO (42 CFR 1004.1b). A gross and flagrant violation entails a violation of an obligation in one or more instances which presents an imminent danger to the health, safety, or well-being of a Medicare beneficiary or places the beneficiary in high-risk situations (42 CFR 1004.1b)

The SuperPRO

In June 1985, HCFA contracted with SysteMetrics, Inc. to evaluate the PRO program. Every 6 months, this organization, also known as the SuperPRO, re-reviews a random sample of 400 medical records from each of the 54 PROs' random sample of reviews (199). The purposes of the SuperPRO's reviews are as follows:

- to validate the determinations made by PROs, specifically on admission review, discharge review, and DRG validations;
- to validate the medical review criteria used by nonphysician reviewers for admission reviews;
- to verify that nonphysicians are properly applying the PROs' criteria for referring cases to physicians for review; and
- to identify quality issues that should have been addressed by the PRO (use of screening criteria) (637).

The SuperPRO submits the reports generated for each PRO to HCFA. Problems identified by the SuperPRO are also submitted to the individual PRO. The PRO may appeal the SuperPRO's findings with additional data or explanations. If PRO appeals do not lead to a reversal of the initial SuperPRO findings, HCFA reviews the SuperPRO findings and initiates appropriate actions to correct any problems. HCFA is responsible for any final determinations (637).

The SuperPRO review process is mostly educational for the PROs. The SuperPRO's record review may detect an aberrant pattern of care not recognized by the PRO's initial review. Thus, PROs are made aware of the types of cases that should be addressed differently.

Similar to the PRO review process, the SuperPRO has a team of chart reviewers that initially evaluates the hospital records (without benefit of the PRO's reviewer findings).[24] A subcontractor has recruited physicians from across the country (providing a representative sampling of medical specialties and geographical regions), and they make the ultimate decisions on the medical necessity of the admission, DRG validation, appropriateness of discharge, and quality of care (83). The SuperPRO uses the same basic screens in its review that each individual PRO used for the initial review.

HCFA conducts its own review of PRO activities through an internal PRO Monitoring Protocol and Tracking System. This system evaluates how well PROs have fulfilled their contractual obligations. If the PRO data reveal that a PRO is not performing adequately,[25] a corrective action plan may be instigated by HCFA regional offices. Deficiencies in areas such as the use of generic screens, physician profiles, or timeliness of reviews may warrant a corrective action plan. Although data from each PRO are collected by HCFA regional offices every 9 months, a final evaluation of a PRO's contractual performance is not conducted until 90 days before the PRO contract's expiration date (429).

---

[24] The records reviewed by the SuperPRO are copies of the hospital records used by the PRO in the initial review process. The PRO must copy the medical records as requested and send them to the SuperPRO.

[25] HCFA regional offices record the frequencies and the percentage of cases for which they disagreed with initial PRO determinations.

Appendix E

Selected Studies Related to the Quality of Medical Care[a]

Study	Period	Funding
FEDERAL STUDIES		
Health Care Financing Administration:		
1. Nonintrusive Outcome Measures: Identification and Validation	9/84 to mid/88	$865,000
2. Impact of the DRG-Based Prospective Payment System on Quality of Care for Hospitalized Medicare Patients	9/85 to 9/88	$3,500,000
3. Prospective Payment Beneficiary Impact Study	Ongoing	Intramural
4. The Impact of the Prospective Payment System on the Quality of Inpatient Care	9/84 to 9/88	$275,689
5. Indexes of Hospital Efficiency and Quality	3/86 to 12/87	$227,097
6. Pilot Study of the Appropriateness of Post-Hospital Care Received by Medicare Beneficiaries	3/86 to 8/87	$1,133,000
7. Learning From and Improving Diagnosis-Related Groups for End-Stage Renal Disease Patients	9/84 to 12/87	$375,500
8. Changes in Post-Hospital Service Use by Medicare Beneficiaries	9/85 to 6/87	$203,600
9. Patient Classification Systems: An Evaluation of the State of the Art	7/87 to 6/89	$1,602,544
10. An Automated, Data-Driven, Case-Mix Adjustment System for Studies of Quality of Care	7/87 to 6/90	$526,948
11. Trends in Patterns of Post-Hospital Service Use and Their Impacts on Outcomes	6/87 to 5/90	$293,922
12. Evaluating Outcomes of Hospital Care Using Claims Data[b]	7/87 To 6/89	$500,000
13. Methods To Improve Case-Mix and Severity of Illness Classification for Use in the Medicare Prospective Payment System	9/85 to 7/88	$1,013,395
14. Impact of the Prospective Payment System on Post-Hospital Care Among Medicare/Medicaid Recipients	8/86 to 6/87	$111,969
15. Impact of the Prospective Payment System on Mortality Rates: Adjustments for Case-Mix Severity	8/86 to 6/87	$125,000
16. Health Status at Discharge Research Project	9/85 to 6/87	$68,000
17. Pneumococcal Pneumonia Immunization in the Baltimore Area	2/88 to 8/88	$24,000
18. National End-Stage Renal Disease Registry	2/88 to 2/93	$6,000,000
19. Study of the End-Stage Renal Disease Program	10/88 to 12/90	In negotiation
20. Strategies for Assessing and Assuring Quality of Care in the Medicare Program	10/87 to 12/89	$1,900,000
21. Mortality Predictors Project	8/87 to 9/88	$600,000
22. Influenza Vaccine Demonstration	10/88 to 9/90[c]	$25,000,000 per year
23. Quality of Care: Selected Issues in Medicaid	9/88 to 9/90	In negotiation
24. Health Care Outcomes by Geographic Area	9/87 to 4/88	Intramural
25. Alternative Outcomes Study	9/87 to 4/88	Intramural
26. A National Program To Improve the Quality of ICU Services	1/88 to 12/90	$770,000
27. Impact of the Prospective Payment System on the Quality of Long-Term Care in Nursing Homes and Home Health Agencies	8/86 to 1/88	$374,011
28. Development, Pilot Testing, and Refinement of Valid Outcome Measures for the Home Care Setting	9/85 to 8/88	$188,766
29. Special Projects for the Monitoring of Quality of Care (PRO Pilot Project)	12/86 to 7/88	Intramural
30. Development of Uniform Clinical Data Set and Screening Algorithms	9/87 to 1/89	Intramural

Study	Period	Funding
National Center for Health Services Research:		
1. Appropriate and Inappropriate Use of Ancillary Services	9/85 to 5/89	$791,864
2. Dental Demand and Public Oral Health	2/86 to 3/88	$172,371
3. Variations in Hospital and Physician Resource Use	2/86 to 1/88	$164,401
4. The Effect of Computerization on the Nursing Process	7/86 to 6/89	$151,400
5. Doctor-Child Communication: Improving Health Outcomes	2/86 to 1/89	$385,107
6. Small Area Analysis of Surgery for Back Pain	2/87 to 6/88	$69,949
7. Research in Trauma Indices	7/86 to 6/89	$217,500
8. Impact of DRGs on Public Home Health Nursing Services	9/86 to 2/89	$454,722
9. Information Retrieval in National Survey	3/86 to 2/88	$137,504
10. AIDS and Other Patients' Use of ICU Technologies	5/88 to 10/89	$85,394
11. Assessment of Coronary Care Unit Use in Different Hospitals	5/87 to 4/89	$454,797
12. Impact of Payment Restrictions on Physician Decisions	9/86 to 8/88	$371,617
13. Comparison of Extracorporeal and Percutaneous Lithotripsy	9/87 to 8/89	$324,816
14. Quality Differences Among Primary Care Practitioners	9/86 to 9/88	$21,550
15. New Techniques for Pretesting Survey Questions	9/87 to 2/89	$212,509
16. Effects of Advance Directives in Medical Care	6/87 to 11/90	$625,191
17. International Collaborative Study of Oral Health Outcomes	6/87 to 5/92	$503,049
18. Variations in Coronary Artery Bypass: Determinants and Clinical Significance	9/86 to 9/88	$21,600
19. Validation of AIS[d] and ISS[e] for Pediatric Trauma	9/86 to 9/88	$79,460
20. Statistical Methods for Longitudinal Health Data	7/87 to 6/90	$161,281
21. Determinants of Inappropriate Hospital Admissions	9/87 to 9/88	$32,509
22. Evaluating Outcomes of Hospital Care Using Claims Data[b]	9/87 to 8/89	$1,342,809
23. Teaching Effects on Outcomes and Costs of Patient Care	7/87 to 6/88	$21,600
24. Technology Assessment: Evaluation of Electronic Fetal Monitoring	7/87 to 6/88	$20,198
25. A Research Agenda on Rural Health	2/87 to 1/88	$190,000
26. Variability in "No Code" Orders Among Very Ill Patients	9/87 to 9/88	$64,544
27. The Effect of WIC[f] on Adolescent Pregnancy Outcomes	9/87 to 9/88	$21,578
28. Diffusion of New Drug Technologies	9/87 to 9/88	$19,647
29. Evaluation of Family Impact of Home Apnea Monitoring	9/87 to 3/89	$21,471
30. Variations in Physician Practice Style and Outcomes of Care	1/88 to 12/90	$539,389
General Accounting Office:		
1. How Effective is the Credentialing and Delineation of Privileges at VA Medical Centers	9/87 to 11/88	Intramural
2. Study of the VA's Infection Control Program	3/87 to 11/88	Intramural
3. Review of the DOD's Health Care Licensure and Credentialing Systems	10/85 to 2/88	Intramural
4. Medicare: Improving Quality of Care Assessment and Assurance	3/86 to 4/88	Intramural
5. Medicare: Improved Patient Outcome Analyses Could Enhance Quality Assessment	1/87 to 6/88	Intramural
Prospective Payment Assessment Commission:		
1. Trends and Concentration of Specialized Procedures Under the Prospective Payment System	10/87 to (work still in progress)	$49,797
2. Assessing Quality Assurance Software Packages	6/87 to 11/87	$14,417
3. Evaluating PRO Activities in the Preadmission Review Process	6/87 to 12/87	$34,442
4. Adjustment Methodologies for Outcome Statistics	9/87 to 1/88	$18,206
NONFEDERAL STUDIES		
American College of Physicians:		
1. Clinical Efficacy Assessment Project	Ongoing	$140,000 per year

Study	Period	Funding
American Medical Review Research Center:		
1. Small Area Analysis of Variation in Utilization and Outcome	10/87 to 6/90	$1,730,000
Cigna Foundation		
1. The Quality of Care Initiatives—to promote the private sector's capacity to measure and monitor the quality of health care	11/87 to 11/88	$150,000
Henry J. Kaiser Foundation:		
1. National Study of Medical Care Outcomes[g]	12/84 to 1989	$3,600,000
2. Improving the Functional Outcomes of Chronic Arthritis Patients by Improving Rheumatological Practice	11/86 to 1991	$443,498
Hospital Research and Educational Trust:[h]		
1. The Quality Measurement Task Force	7/87 to 3/88	$150,000[i]
John A. Hartford Foundation:		
1. Monitoring the Quality of Care in Capitated Systems of Health Care	11/87 to 11/88	$205,481
2. To foster a more rational and uniform approach to data collection among State health care data agencies and others	1/87 to 1989	$250,000
3. Managed Care Development Project and Epidemiological Model for Quality Assessment in an HMO	4/87 to 3/89	$408,000
4. The Value Managed Health Care Purchasing Project	1/88 to 1/90	$200,000
Joint Commission for the Accreditation of Healthcare Organizations:		
1. The "Agenda for Change" (including the development of clinical indicators and risk adjustment methods across all health care settings)	Ongoing	Approx. $2,000,000 per year, from multiple sources
The Pew Charitable Trusts:		
1. National Study of Medical Care Outcomes[g]	1986 through 4 yrs	$1,000,000
2. The Effects of Primary Nursing versus Team Nursing on the Quality and Cost of Inpatient Care	1986 through 3 yrs	$200,000
3. Quality, Case Mix, and the Cost of Hospital Care	2/87 through 2 yrs	$349,000
Robert Wood Johnson Foundation:		
1. To understand the relationship of patient and physician characteristics to the appropriateness of care	8/87 to 1/89	$225,000
2. Testing and Evaluation of a Statewide Quality Indicator Screening System for Hospital Trustees and Physicians	1/88 to 12/90	$368,000
3. Utilization management system to assure quality care	4/86 to 3/88	$272,966
4. Develop an agenda for a national committee for quality assurance for HMOs	1/88 to 6/88	$49,100
5. A cooperative approach to quality and credentialing in rural Wisconsin hospitals	1/88 to 12/90	$113,886

[a] This appendix lists studies that were in progress in April 1988 or had been completed in the preceding 11 months.

[b] This study is funded jointly by HCFA and NCHSR.

[c] Study will continue until 9/92 if it is proven cost-effective.

[d] The Abbreviated Injury Scale.

[e] The Injury Severity Scale.

[f] The Special Supplement Food Program for Women, Infants, and Children.

[g] This study is funded jointly by the Henry J. Kaiser Family Foundation, the Pew Charitable Trusts, the Robert Wood Johnson Foundation, the National Center for Health Services Research, and the National Institute of Mental Health.

[h] This group is an affiliate of the American Hospital Association.

[i] These data are the time period and funding for the first phase of the project.

SOURCE: Office of Technology Assessment, 1988.

References

References

1. Abdellah, F.G., and Levine, E., "Developing a Measure of Patient and Personnel Satisfaction With Nursing Care," *Nursing Research* 5:100-108, 1957.
2. Abdellah, F.G., and Levine, E., *Effect of Nurse Staffing on Satisfaction With Hospital Care*, Hospital Monograph Series No. 4 (Chicago, IL: American Hospital Association, 1958).
3. Abramowitz, S., Cote, A.A., and Berry, E., "Analyzing Patient Satisfaction: A Multianalytic Approach," *Quality Review Bulletin* 13:122-130, 1987.
4. Accreditation Association for Ambulatory Health Care, Inc., *Accreditation Handbook for Ambulatory Health Care, 1987-88 Edition* (Skokie, IL: 1987).
5. Adams, D.F, Fraser, D.B., and Abrams, H.L., "The Complications of Coronary Arteriography," *Circulation* 48(3):609-618, 1973.
6. Adams, E.K., and Zuckerman, S., "Variation in the Growth and Incidence of Medical Malpractice Claims," *Journal of Health Politics, Policy, and Law* 9:475-488, Fall 1984.
7. Aday, L.A., and Andersen, R.A., *Access to Medical Care* (Ann Arbor, MI: Health Administration Press, 1975).
8. Aday, L.A., Andersen, R.A., and Fleming, G.V., *Health Care in the U.S.: Equitable for Whom?* (Beverly Hills, CA: Sage Publishing Co., 1980).
9. Advisory Board for Osteopathic Specialists, "Requirements for Certification: Advisory Board for Osteopathic Specialists and Boards of Certification," Chicago, IL, 1987.
10. Affeldt, J.E., Roberts, J.S., and Walczac, R.M., "Quality Assurance—Its Origin, Status, and Future Direction—A JCAH Perspective," *Evaluation and the Health Professions* 6(2):245-248, June 1983.
11. Ahlgren, L., Associate Director, Department of Research and Development, Joint Commission on the Accreditation of Healthcare Organizations, Chicago, IL, personal communication, Feb. 19, 1987.
12. Aleshire, P., "Eastbay Hospitals Top State Death Rates," *The Tribune* (Oakland, CA), p. D1, Sept. 18, 1986.
13. Alpert, J.J., Kosa, J., Haggerty, R.J., et al., "Attitudes and Satisfactions of Low-Income Families Receiving Comprehensive Pediatric Care," *American Journal of Public Health* 60:499-506, 1970.
14. Altman, S.H., Chairman, Prospective Payment Assessment Commission, Washington, DC, letter to Director of the Health Care Financing Administration, U.S. Department of Health and Human Services, Nov. 13, 1987.
15. American Academy of Pediatrics and the American College of Obstetricians and Gynecologists, *Guidelines for Perinatal Care* (Evanston, IL: 1983).
16. American Board of Internal Medicine, "Policies and Procedures," Philadelphia, PA, 1988.
17. American Board of Medical Specialties, "Compendium of Certified Medical Specialists," Evanston, IL, 1986.
18. American Board of Medical Specialties, *Annual Report & Reference Handbook* (Evanston, IL: 1987).
19. American Board of Medical Specialties, "Medical Specialty Certification and Related Matters," Evanston, IL, 1987.
20. American Board of Medical Specialties, "Self-Designated Boards," Evanston, IL, June 16, 1987.
21. American Board of Medical Specialties, Committee on Study of Evaluation Procedures, "Suggestions on Evaluating Residents for Specialty Boards," Evanston, IL, 1986.
22. American College of Cardiology, "The 17th Bethesda Conference—Adult Cardiology Training," Bethesda, MD, 1986.
23. American College of Emergency Physicians and Emergency Nurses Association, "Emergency Care Guidelines (revised)," *Annals of Emergency Medicine* 15:486-490, Apr. 4, 1986.
24. American College of Physicians, *Guide for the Use of ACP Statements on Clinical Competence* (Philadelphia, PA: March 1987).
25. American College of Surgeons, "Statements on Principles," Chicago, IL, June 1985.
26. American College of Surgeons, *Hospital and Prehospital Resources for Optimal Care of the Injured Patient and Appendices A Through J* (Chicago, IL: 1986).
27. American College of Surgeons, Commission on Cancer, *Cancer Program Manual 1986* (Chicago, IL: 1986).

28. American College of Surgeons, Commission on Cancer, "Hospital Cancer Program Fact Sheet," Chicago, IL, undated.
29. American Hospital Association, *American Hospital Association Guide to the Health Care Field* (Chicago, IL: Joint Commission on the Accreditation of Hospitals, 1986).
30. American Hospital Association, *Medical Malpractice Task Force Report on Tort Reform and Compendium of Professional Liability Early Warning Systems for Health Care Providers* (Chicago, IL: 1986).
31. American Medical Association, *The American Medical Directory* (Chicago, IL: 1986).
32. American Medical Association, "Statutes on Medical Disciplinary Boards," *State Health Legislation Report* 14(3):14-25, 1986.
33. American Medical Association, "AMA Initiative on Quality of Medical Care and Professional Self Regulation," Chicago, IL, June 1986.
34. American Medical Association, "Reports of Council on Medical Service," Chicago, IL, June 1986.
35. American Medical Association, *Physician Characteristics and Distribution: 1986* (Chicago, IL: 1987).
36. American Medical Association, "Report of the Board of Trustees: Categorization of Hospital Emergency Capabilities," Chicago, IL, 1987.
37. American Medical Association, *Seeking Quality Medical Care: What You Should Know* (Chicago, IL: 1987).
38. American Medical Association, Commission on Emergency Medical Services, *Provisional Guidelines for the Optimal Categorization of Hospital Emergency Capabilities* (Chicago, IL: 1982).
39. American Medical Association, Division of Health Policy and Program Evaluation, Department of Health Care Review, "Confronting Regional Variations: The Maine Approach," Chicago, IL, 1986.
40. American Medical Association, Division of Survey and Data Resources, "Intended Use of AMA Physician Masterfile Codes for Self-Designation of Practice Specialty," Chicago, IL, January 1987.
41. American Medical News, "Battle Over Performance Monitoring Shapes Up," *American Medical News*, p. 1, Oct. 16, 1987.
42. American Psychological Association, *Standards for Educational and Psychological Tests* (Washington, DC: American Psychological Association, 1974).
43. Anderson, J.G., and Bartkus, D.E., "Choice of Medical Care: A Behavioral Model of Health and Illness Behavior," *Journal of Health and Social Behavior* 14(4):348-362, 1973.
44. Ashworth, J.W., Executive Director, The Maryland Institute for Emergency Medical Services Systems, University of Maryland, Baltimore, MD, personal communication, Feb. 18, 1988.
45. Association of Community Cancer Centers, "Application for Membership," Rockville, MD, undated.
46. Association of Community Cancer Centers, *Standards* (Rockville, MD: undated).
47. Australian National Blood Pressure Management Committee, "The Australian Therapeutic Trial in Mild Hypertension," *Lancet* 1:1261-1267, 1980.
48. Avant, D., Vice President for Accreditation Surveys, Joint Commission on the Accreditation of Healthcare Organizations, Chicago, IL, personal communication, Dec. 30, 1987, and January 1988.
49. Ayliffe, G.A.J., "Nosocomial Infection: The Irreducible Minimum," *Infection Control* 7:92-95, 1986.
50. Bachman, S.S., Pomerantz, D., and Tell, E., "Making Employers Smart Buyers of Health Care," *Business and Health* 4(9):28-34, 1987.
51. Ball, J.R., "Credentials vs. Privileges: Another Look," *Hospital Privileges and Specialty Medicine* (Chicago, IL: American Board of Medical Specialties, 1986).
52. Ballard, D.J., Strogatz, D.S., Wagner, E.H., et al., "The Edgecombe County High Blood Pressure Control Program: The Process of Medical-Care and Blood Pressure Control," *American Journal of Preventive Medicine* 2(5):278-284, 1986.
53. Banta, H.D., Behney, C.B., and Willems, J.S., *Toward Rational Technology in Medicine* (New York, NY: Springer Publishing Co., 1981).
54. Banton, L., Secretary Peer Assessment Committee, The College of Physicians and Surgeons of Ontario, Toronto, Ontario, personal communication, Dec. 2, 1987.
55. Bargmann, E., and Grove, C., *Surgery in Maryland Hospitals 1979 and 1980: Charges and Deaths* (Washington, DC: Public Citizen Health Research Group, 1982).
56. Bartlett, E.E., Grayson, M., Barker, R., et al., "The Effects of Physician Communications Skills on Patient Satisfaction, Recall, and Adherence," *Journal of Chronic Diseases* 37:755-764, 1984.

57. Batalden, P., Vice President for Medical Care, Hospital Corporation of America, Nashville, TN, personal communication, Aug. 3, 1987.
58. Batalden, P.B., and O'Connor, J.P., *Quality Assurance in Ambulatory Care* (Rockville, MD: Aspen Publications, 1980).
59. Bates, B., *A Guide to Physical Examination and History Taking* (Philadelphia, PA: J.B. Lippincott Co., 1987).
60. Bauer, R.A., "Consumer Behavior as Risk Taking," *Dynamic Marketing for a Changing World*, R.S. Hancock (ed.) (Chicago, IL: American Marketing Association, 1960).
61. Becker, M.H., "Psychosocial Aspects of Health-Related Behavior," *Handbook of Medical Sociology*, H.E. Freeman, S. Levine, and L.G. Reeder (eds.) (Englewood Cliffs, NJ: Prentice-Hall, Inc., 1979).
62. Becker, M.H., "Patient Adherence to Prescribed Therapies," *Medical Care* 23(5):539-555, 1985.
63. Becker, M.H., Maiman, L.A., Kirscht, J.P., et al., "The Health Belief Model and Prediction of Dietary Compliance: A Field Experiment," *Journal of Health and Social Behavior* 18:348-366, December 1977.
64. Belsky, M.S., and Gross, L., *Beyond the Medical Mystique: How To Choose and Use Your Doctor* (New York, NY: Priam Books, 1975).
65. Bennett, D., and Campbell, K., "Most Consumers Show Little Concern in Selecting Health-Services Providers," *Marketing News*, p. 14, Aug. 15, 1986.
66. Bertakis, K.D., "The Communication of Information From Physician to Patient: A Method for Increasing Patient Retention and Satisfaction," *Journal of Family Practice* 5:217-222, 1977.
67. Berwick, D. (Harvard Community Health Plan, Cambridge, MA), and Batalden, P. (Hospital Corporation of America, Nashville, TN), "Toward an Alternative Theory of Quality Improvement," 1987.
68. Berwick, D., and Godfrey, A.B., "Industrial Quality Control and Health Care," Harvard Community Health Plan, Cambridge, MA, August 1986.
69. Berwick, D., Ware, J.E., Jr., Nelson, E., et al., "Patient Judgements of Hospital Quality: Report of a Pilot Study," Harvard Community Health Plan, Cambridge, MA, forthcoming.
70. Bettman, J.R., *An Information Processing Theory of Consumer Choice* (Reading, MA: Addison-Wesley, 1979).
71. Billings, J., and Eddy, E., "Physician Decision Making Limited by Evidence," *Business and Health* 5(1):23-28, 1987.
72. Blendon, R.J., Aiken, L.H., Freeman, H.E., et al., "Uncompensated Care by Hospitals or Public Insurance for the Poor: Does it Make a Difference?" *New England Journal of Medicine* 314(18):1160-1163, 1986.
73. Blue Cross and Blue Shield of Ohio, "Consumer Guide for Patients and Physicians," Cleveland, OH, May 1987.
74. Blumberg, M.S., "Regional Differences in Hospital Use Standardized by Reported Morbidity," *Medical Care* 20:931-944, September 1982.
75. Blumberg, M.S., "At Risk for Hospitalization: Differences by Health Insurance Coverage and Income," *Advances in Health Economics and Health Services Research, Vol. 5*, R.M. Scheffler and L.F. Rossiter (eds.) (Greenwich, CT: JAI Press, 1984).
76. Blumberg, M.S., "Measures of Risk for Short-Term Hospital Days and Part B Covered Charges in the 1977 CMS Aged Sample," Kaiser Foundation Health Plan, Oakland, CA, Apr. 10, 1985.
77. Blumberg, M.S., "Risk-Adjusting Health Care Outcomes: A Methodologic Review," *Medical Care Review* 43:351-393, 1986.
78. Blumberg, M.S., "Comments on HCFA Hospital Death Rate Statistical Outliers," *Health Services Research* 21(6):715-739, February 1987.
79. Blumberg, M.S., "Scholarly Debate: Inter-Area Variations in Age-Adjusted Health Status," *Medical Care* 25(4):340-353, April 1987.
80. Blumberg, M.S., "Maryland Mortality for Non-Elective Surgery: A Prototype RAMO System," Oakland, CA, Kaiser Foundation Health Plan, Inc., May 6, 1987.
81. Blumberg, M.S., "Measuring Surgical Quality in Maryland: A Model," *Health Affairs* 7(1):62-78, 1988.
82. Bodendorf, F.M., Assistant Director, Pennsylvania Health Care Cost Containment Council, Harrisburg, PA, personal communication, Dec. 30, 1987, and Feb. 18, 1988.
83. Booth, P., Branch Chief, Utilization Review Branch, Health Care Financing Administration, U.S. Department of Health and Human Services, Bethesda, MD, personal communication, Feb. 11, 1988, Mar. 2, 1988, and Apr. 4, 1988.
84. Borgiel, A.E.M., "Assessing the Quality of Care in Family Physicians' Practices by the College of Family Physicians of Canada," presentation

at the Institute of Medicine Health Care Technology Forum on Quality of Care and Technology, Washington, DC, May 15, 1987.

85. Borgiel, A.E.M., Chairman, Practice Assessment Committee, College of Family Physicians of Canada, Mississauga, Ontario, personal communication, August 1987, and Sept. 28, 1987.
86. Borgiel, A.E.M., Williams, J.I., Anderson, G.M., et al., "Assessing the Quality of Care in Family Physicians' Practices," *Canadian Family Physician* 31(4):853-862, 1985.
87. Bovbjerg, R., "Medical Malpractice on Trial: Quality of Care Is the Important Standard," *Law and Contemporary Problems* 49(2):321-348, spring 1986.
88. Bovbjerg, R., Senior Research Associate, Urban Institute, Washington, DC, personal communication, Dec. 1, 1987.
89. Braunwald, E., Isselbacher, K.J., Petersdorf, R.G., et al. (eds.), *Harrison's Principles of Internal Medicine* (New York, NY: McGraw Hill Book Co., 1987).
91. Breaden, D.G., Associate Executive Vice President, The Federation of State Medical Boards of the United States, Fort Worth, TX, personal communication, Sept. 30, 1987, Oct. 13, 1987, and Feb. 9, 1988.
90. Breaden, D.G., and Galusha, B.L., "Official 1985 Federation Summary of Reported Disciplinary Actions," *Federation Bulletin* 73:300-305, 1986.
92. Breslau, N., and Mortimer, E.A., "Seeing the Same Doctor: Determinants of Satisfaction With Specialty Care," *Medical Care* 19:741-758, 1981.
93. Brewster, A.C., Jacobs, C.M., and Bradbury, R.C., "Special Report: Classifying Severity of Illness by Using Clinical Findings," *Health Care Financing Review*, Annual Supplement, pp. 107-109, December 1984.
94. Brewster, A.C., Karlin, B.G., Hyde, L.A., et al., "MEDISGRPS: A Clinically Based Approach to Classifying Hospital Patients at Admission," *Inquiry* 22:377, 1985.
95. Britt, M.R., Burke, J.P., Nordquist, A.G., et al., "Infection Control in Small Hospitals: Prevalence Surveys in 18 Institutions," *Journal of the American Medical Association* 236(15):1700-1703, 1976.
96. Britt, M.R., Schleupner, C.J., and Matsumiya, S., "Severity of Underlying Disease as a Predictor of Nosocomial Infection: Utility in the Control of Nosocomial Infection," *Journal of the American Medical Association* 239(11):1047-1051, 1978.
97. Brody, B.A., "HCFA Data Release—Government Abuse or Patient Right?" *Hospital Physician*, pp. 36-40, June 1986.
98. Brook, H., Deputy Director, Office of Medical Review, Health Care Financing Administration, U.S. Department of Health and Human Services, Baltimore, MD, personal communication, Mar. 16, 1988.
99. Brook, R.H., *Quality of Care Assessment: A Comparison of Five Methods of Peer Review* (Washington, DC: U.S. Department of Health, Education, and Welfare, 1973).
100. Brook, R.H., and Appel, F.A., "Quality-of-Care Assessment: Choosing a Method for Peer Review," *New England Journal of Medicine* 288(25):1323-1329, 1973.
101. Brook, R.H., Brutoco, R., and Williams, K.N., "The Relationship Between Medical Malpractice and Quality of Care," *Duke Law Journal* 1975:1197-1231, 1975.
102. Brook, R.H., Fink, A., Kosecoff, J., et al., "Educating Physicians and Treating Patients in the Ambulatory Setting: Where Are We Going and How Will We Know When We Arrive?" *Annals of Internal Medicine* 107:392-398, 1987.
103. Brook, R.H., and Lohr, K.N., "Efficacy, Effectiveness, Variations and Quality: Boundary-Crossing Research," *Medical Care* 23(5):710-722, May 1985.
104. Brook, R.H., and Lohr, K.N., "Monitoring Quality of Care in the Medicare Program: Two Proposed Systems," *Journal of the American Medical Association* 258(21):3138-3141, 1987.
105. Brook, R.H., and Williams, K.N., "Quality of Health Care for the Disadvantaged," *Journal of Community Health* 1(2):132-156, 1975
106. Budetti, P., McManus, P., Barrand, N., et al., *The Costs and Effectiveness of Neonatal Intensive Care* (Health Technology Case Study #10), prepared for the Office of Technology Assessment, U.S. Congress, OTA-BP-H-9 (Washington, DC: U.S. Government Printing Office, December 1981); also available as PB 82-101 411 (Springfield, VA: National Technical Information Service, August 1981).
107. Bunker, J.P., and Brown, B.W., "The Physician-Patient as an Informed Consumer of Surgical Services," *New England Journal of Medicine* 290(19):1051-1055, 1974.
108. Bunker, J.P., Forrest, W.H., Jr., Mosteller, F., et al., *The National Halothane Study* (Washington, DC: U.S. Government Printing Office, 1969).
109. Burda, D., "Physician Data Base Spurs Legal, Ethical Debate," *Hospitals* 61(2):56, Jan. 20, 1987.

110. Burg, F.D., and Lloyd, J.S., "Definitions of Competence: A Conceptual Framework," *Evaluating the Skills of Medical Specialists* (Chicago, IL: American Board of Medical Specialties, 1983).
111. Cahill, N.E., "Developing Law on Professional Standards and Peer Review in Quality Assessment Activities," contractor document prepared for the Office of Technology Assessment, U.S. Congress, Washington, DC, December 1987.
112. California Blue Shield, "Bypass Surgery Survival Odds Much Better at Some Hospitals: Blue Shield Study Finds Lives and Dollars Could Be Saved," press release, Rencho Cordova, CA, May 21, 1986.
113. California Department of Health Services, *Licensing and Certification, Facility Listing* (Sacramento, CA: 1987).
114. California Medical Association and California Hospital Association, *Report on the Medical Insurance Feasibility Study* (San Francisco, CA: Sutter Publications, Inc., August 1977).
115. California Medical Review, Inc., "CMRI Releases Medicare Data for California Hospitals," CMRI press release, San Francisco, CA, Aug. 18, 1986.
116. California Medical Review, Inc. "California Medical Review, Inc. (CMRI) Releases DRG-Specific Medicare Data on State Hospitals: Executive Summary," San Francisco, CA, 1987.
117. California Medical Review, Inc., San Francisco, CA, *Premature Discharge Study* prepared for the Health Care Financing Administration, U.S. Department of Health and Human Services, undated.
118. Cancer Letter, "CCOP, Another 'Jewel,' Into Second Life Following Bitter Recompetition," *Cancer Letter* 14(8):3-8, 1988.
119. Cancila, C., "JCAH to Review Clinical Outcomes," *American Medical News*, p. 1, Sept. 19, 1986.
120. Carmel, S., "Satisfaction With Hospitalization: A Comparative Analysis of Three Types of Services," *Social Science and Medicine* 21:1243-1249, 1985.
121. Carney, S.L., and Mitchell, K.R., "Satisfaction of Patients With Medical Students' Clinical Skills," *Journal of Medical Education* 61:374-379, 1986.
122. Carper, J., *Health Care, U.S.A.* (New York, NY: Prentice Hall Press, 1987).
123. Carter, W.B., Inui, T., Kukull, W., et al., "Outcome-Based Doctor-Patient Interaction Analysis, Part II: Identifying Effective Provider and Patient Behaviors," *Medical Care* 20:550-566, 1982.
124. Casarreal, K.M., Mills, J.I., and Plant, M.A., "Improving Service Through Patient Surveys in a Multihospital Organization," *Hospital and Health Services Administration* 31:41-52, 1986.
125. *Cassim v. Bowen*, 824 F.2d 791 (9th Cir. 1987).
126. Center for Medical Consumers, "Where To Go for Coronary Bypass Surgery: Special Report," Center for Medical Consumers, New York, NY, 1986.
127. Champion, H., Chief Director, Critical Surgical Care Services, Medstar, Washington Hospital Center, Washington, DC, personal communication, March 1988.
127b. Champion, H. (Chief Director, Critical Surgical Care Services, Medstar, Washington Hospital Center), and Teter, H. (Attorney at Law, Bricker & Eckler), "State Laws and Regulations," Washington, DC, 1988.
128. Chang, B.L., Uman, G.C., Linn, L.S., et al., "The Effect of Systematically Varying Components of Nursing Care on Satisfaction in Elderly Ambulatory Women," *Western Journal of Nursing Research* 6:367-379, 1984.
129. Charles, J.G., "Using Informed Choice To Combat Health Costs," *Business and Health* 4(9):36-38, 1987.
130. Chase, R.A., and Burg, F.D., "Reexamination/Recertification," *Archives of Surgery* 112(1):19-25, 1977.
131. Chassin, M.R., Kosecoff, J., Park, R.E., et al., "Does Inappropriate Use Explain Geographic Variations in the Use of Health Care Services?" *Journal of the American Medical Association* 258(18):2533-2537, 1987.
132. Chenoweth, J., "Hospital Associations Must Prepare Members for PRO Data Release," *Baselines: A Monthly Newsletter* 1:1-2, 1985.
133. Christoffel, T., and Loewenthal, M., "Evaluating the Quality of Ambulatory Care: A Review of Emerging Methods," *Medical Care* 15:877-897, 1977.
134. Chu, J., Diehr, P., Feigl, P., et al., "The Effect of Age on the Care of Women With Breast Cancer in Community Hospitals," *Journal of Gerontology* 42:185, 1987.
135. Clemmer, T., Orme, J., Thomas, F., et al., "Outcome of Critically Injured Patients Treated at Level I Trauma Centers Versus Full-Service Community Hospitals," *Critical Care Medicine* 13(10):861-863, 1985.
136. Coale, J., Director, Department of Corporate Relations, Joint Commission on the Accredita-

tion of Healthcare Organizations, Chicago, IL, personal communication, Feb. 19, 1987.

137. Cohen, J., *Statistical Power Analysis for the Behavioral Sciences*, revised edition (New York, NY: Academic Press, 1977).

138. Colburn, D., "Shopping for Hospital Care," *Washington Post*, p. D7, Nov. 14, 1987.

139. College of Physicians and Surgeons of Ontario, *Peer Assessment—Annual Report* (Toronto, Ontario: 1987).

140. Colorado Health Data Commission, "Concerning the Measurement and Reporting of Hospital Inpatient Severity and Morbidity," Regulation 87-3, Denver, CO, Jan. 12, 1988.

141. Commission on Professional and Hospital Activities, *Risk-Adjusted Hospital Mortality Norms 1986 Workbook* (Ann Arbor, MI: 1987).

142. Committee on Perinatal Health, *Toward Improving the Outcome of Pregnancy* (White Plains, NY: National Foundation/March of Dimes, 1977).

143. Committee on Trauma, "Hospital and Prehospital Resources for Optimal Care of the Injured Patient," *Bulletin of the American College of Surgeons* 68:11-18, October 1983.

144. Comstock, L.M., Hooper, E.M., Goodwin, J.M., et al., "Physician Behaviors That Correlate With Patient Satisfaction," *Journal of Medical Education* 57:105-112, 1982.

145. Connell, C.M., and Crawford, C.O., "How People Obtain Their Health Information—A Survey in Two Pennsylvania Counties," *Public Health Reports* 103(2):189-195, 1988.

146. Conrad, F., Director, Office of Quality Assurance, Veterans Administration, Washington, DC, personal communication, Feb. 16, 1988.

147. Cook, T., and Campbell, D.T., *Quasi-Experimentation: Design and Analysis Issues for Field Settings* (Chicago, IL: Rand McNally, 1979).

148. Cooper, H.M., *The Integrative Research Review: A Systematic Approach* (Beverly Hills, CA: Sage Publishing Co., 1984).

149. Cope, D.W., Linn, L.S., Leake, B.D., et al., "Modification of Residents' Behavior by Preceptor Feedback of Patient Satisfaction," *Journal of General Internal Medicine* 1:394-398, 1986.

150. Corah, N.L., O'Shea, R.M., Pace, L.F., et al., "Development of a Patient Measure of Satisfaction With the Dentist: The Dental Visit Satisfaction Scale," *Journal of Behavioral Medicine* 7:367-373, 1984.

151. Cornacchia, H.J., and Barrett, S., *Consumer Health: A Guide to Intelligent Decisions* (St. Louis, MO: C.V. Mosby, 1980).

152. Couch, N.P., Tilney, N.L., Rayner, A.A., et al., "The High Costs of Low-Frequency Events," *New England Journal of Medicine* 305:634-637, 1981.

153. Cowley, R.A. (ed.), *Trauma Care, Vol. 1: Surgical Management* (Philadelphia, PA: J.B. Lippincott Co., 1986).

154. Craddick, J.W., *Medical Management Analysis Series, Vol. II: Improving Quality and Resource Management Through Medical Management Analysis* (Rockville, MD: Medical Management Analysis International, Inc., 1987).

155. Crile, G., "How To Keep Down the Risk and Cost of Surgery," *Inquiry* 18:99-101, 1981.

156. Cronbach, L., *Essentials of Psychological Testing*, 4th edition (New York, NY: Harper and Row, 1984).

157. Crowley, A.E., Etzel, S.I., Peterson, E.S., et al., "Undergraduate Education," *Journal of the American Medical Association* 258(8):1013-1020, 1987.

158. Dans, P.E., Weiner, J.P., and Otter, S.E., "Peer Review Organizations: Promises and Pitfalls," *New England Journal of Medicine* 313(18):1131-1137, 1985.

159. Danzon, P., *Medical Malpractice: Theory, Evidence, and Public Policy* (Cambridge, MA: Harvard University Press, 1985).

160. Daschner, F., Nadjem, H., Langmaack, H., et al., "Surveillance, Prevention and Control of Hospital-Acquired Infections," *Infection* 6(6):261-265, 1978.

161. Davies, A.R., and Ware, J.E., Jr., "Involving Consumers in Quality of Care Assessment: Do They Provide Valid Information?" *Health Affairs* 7(1):33-48, Spring 1988.

162. Davies, A.R., Ware, J.E., Jr., Brook, R.H., et al., "Consumer Acceptance of Prepaid and Fee-for-Service Care: Results From a Randomized Controlled Trial," *Health Services Research* 21:429-452, 1986.

163. Davies, A.R., Ware, J.E., Jr., Brook, R.H., et al., *Consumer Acceptance of Prepaid and Fee-for-Service Care: Results From a Randomized Controlled Trial*, R-3219-HHS (Santa Monica, CA: Rand Corp., in press).

164. Dehn, T.G., President of the American Medical Peer Review Association, testimony at hearing on the Medicare Quality Assurance Process before the Subcommittee on Intergovernmental Relations and Human Resources, Committee on Government Operations, House of Representatives, U.S. Congress, Washington, DC, Oct. 20, 1987.

165. Demlo, L.K., "Assuring Quality of Health Care: An Overview," *Evaluation and the Health Professions* 6(2):161-196, 1983.

166. Demlo, L.K., and Campbell, P.M., "Improving Hospital Discharge Data: Lessons From the National Hospital Discharge Survey," *Medical Care* 19(10):1030-1040, 1981.

167. Demlo, L.K., Campbell, P.M., and Brown, S.S., "Reliability of Information Abstracted From Patients' Medical Records," *Medical Care* 16:995-1005, December 1978.

168. Derbyshire, R.C., "Obstacles to Enforcement of Discipline," *Hospital Practice* 18(10):251-262, 1983.

169. Derbyshire, R.C., "The Incompetent Physician," *Hospital Practice* 18(11):30-50, 1983.

170. DesHarnais, S., Research Scientist, Research Services, Commission on Professional and Hospital Activities, Ann Arbor, MI, personal communication, Feb. 11, 1988.

171. DesHarnais, S., Chesney, J., and Wroblewski, R., "Using Data To Evaluate Performance," *Michigan Hospitals*, pp. 21-26, October 1987.

172. DesHarnais, S., Chesney, J., and Fleming, S.T., "Should DRG Assignment Be Based on Age?" *Medical Care* 26:124-131, 1988.

173. DesHarnais, S., Chesney, J., Wroblewski, R., et al., "The Risk-Adjusted Mortality Index: A New Measure of Hospital Performance," *Medical Care*, in press, 1988.

174. Deuschle, J. M., Alvarez, B., Logsdon, D.N., et al., "Physician Performance in a Prepaid Health Plan: Results of the Peer Review Program of the Health Insurance Plan of Greater New York," *Medical Care* 20(2):127-142, 1982.

175. Diamond, S.S., "Order in the Court: Consistency in Criminal Court Decisions," *The Master Lecture Series, Vol. II: Psychology and the Law*, C.J. Scheirer and B.L. Hammonds (eds.) (Washington, DC: American Psychological Association, 1983).

176. Dickinson, J.C., and Gehlbach S.H., "Process and Outcome: Lack of Correlation in a Primary Care Model," *Journal of Family Practice* 7(3):557-562, 1978.

177. Dietrich, A.J., and Marton, K.I., "Does Continuous Care From a Physician Make a Difference?" *Journal of Family Practice* 15(5):929-937, 1982.

178. Dietz, P., Program Administrator, Commission on Emergency Medical Services, American Medical Association, Chicago, IL, personal communication, Jan. 28, 1988.

179. DiFazio, J.E., Associate Director, Association of Community Cancer Centers, Rockville, MD, personal communication, Jan. 7, 1988.

180. DiMatteo, M.R., Hays, R.D., and Prince, L.M., "Relationship of Physicians' Nonverbal Communication Skill to Patient Satisfaction, Appointment Noncompliance, and Physician Workload," *Health Psychology* 5:581-594, 1986.

181. DiMatteo, M.R., Taranta, A., Friedman, H.S., et al., "Predicting Patient Satisfaction From Physicians' Nonverbal Communication Skills," *Medical Care* 18:376-387, 1980.

182. Dixon, R.E., "Costs of Nosocomial Infections and Benefits of Infection Control Programs," *Prevention and Control of Nosocomial Infections*, R.P. Wenzel (ed.) (Baltimore, MD: Williams & Wilkins, 1987).

183. Donabedian, A., *Explorations in Quality Assessment and Monitoring, Vol. I: The Definition of Quality and Approaches to its Assessment* (Ann Arbor, MI: Health Administration Press, 1980).

184. Donabedian, A., *Explorations in Quality Assessment and Monitoring, Vol. II: The Criteria and Standards of Quality* (Ann Arbor, MI: Health Administration Press, 1982).

185. Donabedian, A., *Explorations in Quality Assessment and Monitoring, Vol. III: The Methods and Findings of Quality Assessment and Monitoring: An Illustrated Analysis* (Ann Arbor, MI: Health Administration Press, 1985).

186. Donabedian, A., "A Primer on Quality Assurance and Monitoring in Medical Care," prepared for the Conference on Law Practice Quality Evaluation, American Law Institute and the American Bar Association, Williamsburg, VA, Sept. 10-12, 1987.

187. Dougherty, D., "Hospital Mortality Rates as a Quality Indicator," working paper, Office of Technology Assessment, U.S. Congress, Washington, DC, 1988.

188. Dove, H.G., and Schneider, K.C., "The Application of Prognostic Criteria in Outpatient Quality Assessment: Control of Hypertension," *Journal of Medical Systems* 4(1):45-58, 1980.

189. Dubois, R.W., Brook, R.H., and Rogers, W.H., "Adjusted Hospital Death Rates: A Potential Screen for Quality of Medical Care," *American Journal of Public Health* 77(9):1162-1166, 1987.

190. Dubois, R.W., Rogers, W.H., Moxley, J.H., et al., "Hospital Inpatient Mortality: Is It a Predic-

tor of Quality?" *New England Journal of Medicine* 317(26):1674-1680, 1987.

191. "Dying for Health Care," television show, WETA, Washington, DC, 1987.
192. Eggers, P.W., "Prospective Payment System and Quality: Early Results and Research Strategy," *Health Care Financing Review*, Annual Supplement, pp. 29-37, December 1987.
193. Eggers, P.W., Office of Research, Health Care Financing Administration, U.S. Department of Health and Human Services, Baltimore, MD, personal communication, February 1988.
194. Ehrenhaft, P.M., *Neonatal Intensive Care for Low Birthweight Infants: Costs and Effectiveness* (Health Technology Case Study #38), Office of Technology Assessment, U.S. Congress, OTA-HCS-38 (Washington, DC: U.S. Government Printing Office, December 1987).
195. Ehrlich, J., Morehead, M.A., and Trussell, R.E., *The Quantity, Quality, and Costs of Medical and Hospital Care Secured by a Sample of Teamster Families in the New York Area* (New York, NY: Columbia University School of Public Health and Administrative Medicine, 1961).
196. Eisen, S.A., Woodward, R.S., Miller, D., et al., "The Effect of Medication Compliance on the Control of Hypertension," *Journal of General Internal Medicine* 2:298-305, 1987.
197. *Elam v. College Park Hospital* 132 Cal. App. 3d 333, 183 Cal. Rptr., 156 (1982).
198. Engel, J.F., Blackwell, R.D., and Miniard, P.W., *Consumer Behavior* (Chicago, IL: Dryden Press, 1986).
199. Ernst & Whinney, *The Impact of PROs on Hospitals*, E&W No.J58693 (Cleveland, OH: 1987).
200. Estes, W.K, *Handbook of Learning and Cognitive Processes, Vol. 4: Attention and Memory* (New York, NY: McGraw-Hill, 1959).
201. Evans, R.G., *Strained Mercy: Canada's Health Care System* (Toronto, Ontario: Butterworth & Co. (Canada) Ltd., 1984).
202. Faden, R.R., and Beauchamp, T.L., *A History and Theory of Informed Consent* (New York, NY: Oxford University Press, 1986).
203. Farber, B.F., Kaiser, D.L., and Wenzel, R.P., "Relation Between Surgical Volume and Incidence of Postoperative Wound Infection," *New England Journal of Medicine* 305(4):200-204, 1981.
204. Farrell, K., "Alternative Health Plans: Informing the Medicare Consumer," paper presented at the American Public Health Association Meeting, Las Vegas, NV, September 1986.
205. Federation of State Medical Boards of the United States, "Physician Disciplinary Data Bank," Fort Worth, TX, 1986.
206. Federation of State Medical Boards of the United States, *Exchange/ Section 1: Flex and M.D. Licensing Requirements* (Fort Worth, TX: 1987).
207. Feeny, D., Guyatt, G., and Tugwell, P., *Health Care Technology: Effectiveness, Efficiency, and Public Policy* (Montreal, Quebec: Institute for Research on Public Policy, 1986).
208. Feinstein, R.J., "The Ethics of Professional Regulation," *New England Journal of Medicine* 312(12):801-804, 1985.
209. Fink, A., Brook, R.H., and Yano, E.M., "The Uses and Misuses of Mortality Data: A Literature Review," prepared for the American Medical Review Research Center Executive Training Program on Peer Review Outcome Data, Washington, DC, Apr. 21, 1987.
210. Fishbein, M., and Ajzen, I., *Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research* (Reading, MA: Addison-Wesley, 1975).
211. Fletcher, S.W., Deliakis, J., Schoch, W.A., et al., "Predicting Blood Pressure Control in Hypertensive Patients: An Approach to Quality of Care Assessment," *Medical Care* 17(3):285-292, 1979.
212. Flexner, W.A., *Marketing Research and Health Care: Concepts and Techniques Applied to Abortion Service Design* (New York, NY: Neal Watson Academic Publications, 1978).
213. Flood, A.B., Associate Professor, Medical Humanities and Social Sciences Program, College of Medicine, University of Illinois, Urbana IL, personal communication, Sept. 17, 1987.
214. Flood, A.B., Ewy, W., Scott, W.R., et al., "The Relation Between the Intensity and the Duration of Medical Services and Outcomes for Hospitalized Patients," *Hospital Structure and Performance*, A.B. Flood and W.R. Scott (eds.) (Baltimore, MD: Johns Hopkins University Press, 1987).
215. Flood, A.B., and Scott, W.R., *Hospital Structure and Performance* (Baltimore, MD: Johns Hopkins University Press, 1987).
216. Flood, A.B., and Scott, W.R., "Professional Power and Professional Effectiveness: The Power of the Surgical Staff and the Quality of Surgical Care in Hospitals," *Hospital Structure and Performance*, A.B. Flood and W.R. Scott (eds.) (Baltimore, MD: Johns Hopkins University Press, 1987).

217. Flood, A.B., Scott, W.R., and Ewy, W., "Does Practice Make Perfect? Part I: The Relation Between Hospital Volume and Outcomes for Selected Diagnostic Categories," *Medical Care* 22(2):98-114, 1984.
218. Flood, A.B., Scott, W.R., and Ewy, W., "Reply to David Dranove," letter to the editor, *Medical Care* 22(10):967-969, 1984.
219. Flood, A.B., Scott, W.R., and Ewy, W., "Hospital Characteristics and Hospital Performance," *Hospital Structure and Performance*, A.B. Flood and W.R. Scott (eds.) (Baltimore, MD: Johns Hopkins University Press, 1987).
220. Flood, A.B., Scott. W.R., Ewy, W., et al., "Effectiveness in Professional Organizations: The Impact of Surgeons and Surgical Staff Organizations on the Quality of Care in Hospitals," *Health Services Research* 17(4):341-366, 1982.
221. Flood, A.B., Scott, W.R., Ewy, W., et al., "Effectiveness in Professional Organizations: The Impact of Surgeons and Surgical Staff Oganizations on the Quality of Care," *Hospital Structure and Performance*, A.B. Flood and W.R. Scott (eds.) (Baltimore, MD: Johns Hopkins University Press, 1987).
222. Flood, A.B., Scott, W.R., Ewy, W., et al., "Tests for Differences Among Hospitals in the Quality of Surgical Outcomes and Service Intensity," *Hospital Structure and Performance*, A.B. Flood and W.R. Scott (eds.) (Baltimore, MD: Johns Hopkins University Press, 1987).
223. Forrest, W.H., Jr., Brown, B.W., Jr., Scott, W.R., et al., *Studies of the Determinants of Service Intensity in the Medical Care Sector*, report to the National Center for Health Services Research, PB 287-365 (Springfield, VA: National Technical Information Service, 1977).
224. Fottler, M.D., Slovensky, D.J., and Rogers, S.J., "Public Release of Hospital Specific Death Rates: Guidelines for Health Care Executives," *Hospital and Health Services Administration* 32:343-356, 1987.
225. Fowler, F.J., *Applied Social Research Methods Series, Vol. 1: Survey Research Methods* (Beverly Hills, CA: Sage Publishing Co., 1984).
226. Fowles, J., Vice-President, Research and Development, Park Nicollet Medical Foundation, Minneapolis, MN, personal communication, February 1988.
227. Fowles, J., Bunker, J.P., Oda, M., et al., "Surgery Data Yield Quality Indicators," *Business and Health* 4(8):44-46, 1987.
228. Franklin, F.J., and McLemore, S.D., "A Scale for Measuring Attitudes Toward Student Health Services," *Journal of Psychology* 66(1):143-147, 1967.
229. Franklin, F.J., and McLemore, S.D., "Factors Affecting the Choice of Medical Care Among University Students," *Journal of Health and Social Behavior* 11:311-319, 1970.
230. Freeman, J., and McGowan, J.E., Jr., "Methodologic Issues in Hospital Epidemiology, Part I: Rates, Case Finding and Interpretation," *Review of Infectious Diseases* 3:658-667, 1981.
231. Freeman, J., and McGowan, J.E., Jr., "Methodologic Issues in Hospital Epidemiology, Part III: Investigating the Modifying Effects of Time and Severity of Underlying Illnesses on Estimates of Cost of Nosocomial Infection," *Review of Infectious Diseases* 6(3):285-300, 1984.
232. Freeman, J., and McGowan, J.E., Jr., "Risk Factors for Nosocomial Infections," *Journal of Infectious Diseases* 138:811-819, 1978.
233. Freeman, J., Rosner, B.A., and McGowan, J.E., Jr., "Adverse Effects of Nosocomial Infections," *Journal of Infectious Diseases* 140:732-740, 1979.
234. Freidson, E., *Patients' Views of Medical Practice* (New York, NY: Russell Sage Foundation, 1961).
235. French, K., "Methodologic Considerations in Hospital Patient Opinion Surveys," *International Journal of Nursing Studies* 18(1):7-32, 1981.
236. Friedman, B., Associate Director, Center for Health Services and Policy Research, Northwestern University, Evanston, IL, personal communication, Oct. 26, 1987.
237. Friedman, B., and Shortell, S., "Where Are the Tradeoffs of Cost vs. Patient Outcome in Hospital Care?" Center for Health Services and Policy Research and J.L. Kellogg Graduate School of Management, Northwestern University, Evanston, IL, 1987.
238. Fry, E., Joint Commission on Accreditation of Healthcare Organizations, Chicago, IL, personal communication, July 6, 1987.
239. Galusha, B.L., Executive Vice President, Federation of State Medical Boards of the United States, Fort Worth, TX, personal communication, Sept. 22, 1987.
240. Galusha, B.L., and Breadon, D.G., "Official 1986 Federation Summary of Reported Disciplinary Actions," *Federation Bulletin* 75(2):41-46, 1988.
241. Gardner, "Putting Health Care Decisions in the Hands of Consumers," *Business and Health* 3(10):22-26, 1986.
242. Garg, M.L., Mulligan, J.L., Werner, A.G., et al., "Physician Specialty, Quality and Cost of Inpatient Care," *Social Science and Medicine* 13C:187-190, 1979.

243. General Mills, Inc., "Family Health in an Era of Stress: The General Mills Family Report 1978-1979," Minneapolis, MN: 1979.

244. Georgetown University Hospital, "How To Select the Right Doctor," Washington, DC, April 1987.

245. Georgopolous, B.S., "Organization Structure and the Performance of Hospital Emergency Services," *Annals of Emergency Medicine* 14:677-684, July 7, 1985.

246. Georgopolous, B.S., Research Scientist, Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor, MI, personal communication, Jan. 20, 1988.

247. Gerbert, B., and Hargreaves, W.A., "Measuring Physician Behavior," *Medical Care* 24(9):838-847, 1986.

248. Gerbert, B., Stone, G., Stulbarg, M., et al., "Agreement Among Physician Assessment Methods: Searching for the Truth Among Fallible Methods," *Medical Care*, in press.

249. Giebelhaus, B., Office of Legislation and Policy, Health Care Financing Administration, U.S. Department of Health and Human Services, Washington, DC, personal communication, Feb. 23, 1988, and Mar. 30, 1988.

250. Given, B., Given, C.W., and Simoni, L.E., "Relationships of Processes of Care to Patient Outcomes," *Nursing Research* 28(2):85-93, 1979.

251. Givens, C.D., and Wenzel, R.P., "Catheter-Associated Urinary Tract Infections in Surgical Patients: A Controlled Study on the Excess Morbidity and Costs," *Journal of Urology* 124:646-648, 1980.

252. Glanz, K., "Nutrition Education for Risk Factor Reduction and Patient Education: A Review," *Preventive Medicine* 14(6):721-752, 1985.

253. Glanz, K., and Rudd, J., "Effects of Quality-of-Care Information on Consumer Choice of Physicians and Hospitals," contractor document prepared for the Office of Technology Assessment, U.S. Congress, Washington, DC, December 1987.

254. Glass, G., "Primary, Secondary, and Meta-Analysis of Research," *Educational Researcher* 5:3-8, 1976.

255. Glassman, M., and Glassman, N., "A Marketing Analysis of Physician Selection and Patient Satisfaction," *Journal of Health Care Marketing* 1(4):25-31, 1981.

256. Goldbeck, W.B., "Quality of Care: a Personal Perspective on the Employer's Obligation and Opportunity," draft discussion paper, May 1987, for the Pew Corporate Fellows Conference, Boston, MA, June 10-12, 1987.

257. Goldberg, H.I., and Dietrich, A.J., "The Continuity of Care Provided to Primary Care Patients," *Medical Care* 23(1):63-73, 1985.

258. Goldschmidt, P.G., "The Appropriate Organizational Loci for Constructing Indicators of the Quality of Hospitals and Physicians and for Evaluating the Validity of Those Indicators," contractor document prepared for the Office of Technology Assessment, U.S. Congress, Washington, DC, December 1987.

259. Goss, M.E.W., and Reed, J.I., "Evaluating the Quality of Hospital Care Through Severity-Adjusted Death Rates: Some Pitfalls," *Medical Care* 12:202, 1974.

260. Grad, F.P., and Marti, N., *Physicians' Licensure and Discipline* (Dobbs Ferry, NY: Oceana Publications, 1979).

261. Green, J.W., and Wenzel, R.P., "Postoperative Wound Infection: A Controlled Study of the Increased Duration of Hospital Stay and Direct Cost of Hospitalization," *Annals of Surgery* 185(3):264-268, 1977.

262. Green, L.W., Kreuter, M.W., Deeds, S.G., et al., *Health Education Planning: A Diagnostic Approach* (Palo Alto, CA: Mayfield Publishing Co., 1980).

263. Green, M.S., Rubinstein, E., and Amit, P., "Estimating the Effects of Nosocomial Infections on the Length of Hospitalization," *Journal of Infectious Diseases* 145:667-672, 1982.

264. Greenfield, S., Kaplan, S., and Ware, J.E., Jr., "Expanding Patient Involvement in Care: The Effects on Patient Outcomes," *Annals of Internal Medicine* 102:520-528, 1985.

265. Greenley, J.R., Schulz, R., Nam, S.H., et al., "Patient Satisfaction With Psychiatric Inpatient Care: Issues in Measurement and Application," *Research in Community Mental Health* 4:303-319, 1985.

266. Griffin, N., American Medical Association, Chicago, IL, personal communication, Feb. 19, 1987.

267. Grimaldi, P.L., and Micheletti, J.A., "Implementation of the Peer Review Organization Program," *Quality Review Bulletin* 12:340-346, 1984.

268. Haines, G.H., "Process Models of Consumer Decision Making," *Buyer/Consumer Information Processing*, G.D. Hughes and M.L. Ray (eds.) (Chapel Hill, NC: University of North Carolina Press, 1974).

269. Haley, R.W., Culver, D.H., White, J.W., et al., "The Nationwide Noscomial Infection Rate: A New Need for Vital Statistics," *American Journal of Epidemiology* 121:159-167, 1985.
270. Haley, R.W., Culver, D.H., Morgan, W.M., et al., "Increased Recognition of Infectious Diseases in U.S. Hospitals Through Increased Use of Diagnostic Tests, 1970-1976," *American Journal of Epidemiology* 121(2):168-181, 1985.
271. Haley, R.W., Culver, D.H., Morgan, W.M., et al., "Identifying Patients at High Risk of Surgical Wound Infection: A Simple Multivariate Index of Patient Susceptibility and Wound Contamination," *American Journal of Epidemiology* 121(2):206-215, 1985.
272. Haley, R.W., Culver, D.H., White, J.W., et al., "The Efficacy of Infection Surveillance and Control Programs in Preventing Nosocomial Infections in U.S. Hospitals," *American Journal of Epidemiology* 121(2):182-205, 1985.
273. Haley, R.W., Hooton, T.M., Culver, D.H., et al., "Nosocomial Infections in U.S. Hospitals, 1976-1976: Estimated Frequency by Selected Characteristics of Patients," *American Journal of Medicine* 70:947-959, 1981.
274. Haley, R.W., Quade, D., Freeman, H.E., et al., "The SENIC Project: Summary of Study Design," *American Journal of Epidemiology* 111:472-485, 1980.
275. Haley, R.W., Schaberg, D.R., McClish, D.K., et al., "The Accuracy of Retrospective Chart Review in Measuring Nosocomial Infection Rates," *American Journal of Epidemiology* 111(5):516-533, 1980.
276. *Hall v. Bowen*, No. 86-2380 (8th Cir. 1987).
277. Hamilton, J., "The Best Hospitals for Every Ailment," *Family Circle*, pp.36-40, 133-152, Oct. 20, 1987.
278. Hannan, E.L., Director, Bureau of Health Care Research and Information Services, State of New York Department of Health, Albany, NY, personal communications, October-December 1987.
279. Hannan, E.L., and Yazici, A., "A Critique of the 1987 HCFA Mortality Study Based on New York Data," Office of Health Systems Management, New York State Department of Health, Albany, NY, January 1988.
280. Hansen, H.R., General Counsel, Group Health Association of America, Inc., "Medical Licensure and Consumer Protection: An Analysis and Evaluation of State Licensure," Washington, DC, November 1962.
281. Hanson, F., "Hemispheral Lateralization: Implications for Understanding Consumer Behavior," *Journal of Consumer Research* 8:23-26, June 1981.
282. Hare, R.L., "Let Me Count The Ways," *The Internist* 3:11-13, 1981.
283. Harvard Medical Practice Study Group, "Medical Care and Medical Injuries in the State of New York: A Pilot Study," Cambridge, MA, April 1987.
284. Hastings, G.E., Sonneborn, R., Lee, G.H., et al., "Peer Review Checklist: Reproducibility and Validity of a Method for Evaluating the Quality of Ambulatory Care," *American Journal of Public Health* 70(3):22-28, 1980.
285. Haug, M., and Lavin, B., *Consumerism in Medicine: Challenging Physician Authority* (Beverly Hills, CA: Sage Publishing Co., 1983).
286. Haynes, R.B., Gibson, E.S., Taylor, D.W., et al., "Process Versus Outcome in Hypertension: A Positive Result," *Circulation* 65(1):28-33, 1980.
287. Hays, R.D., and Ware, J.E., Jr., "My Medical Care Is Better Than Yours: Social Desirability and Patient Satisfaction Ratings," *Medical Care* 24(6):519-525, 1986.
288. Headen, K., Program Analyst, Health Care Financing Administration, U.S. Department of Health and Human Services, Baltimore, MD, personal communication, Oct. 16, 1987.
289. Health Technology, "Credentialing for Technology Use," *Health Technology* 1(1):3-10, 1987.
290. Health Technology, "Quality of Care and Technology: How Technology Use Can Have an Important Place in Quality Assurance Programs," *Health Technology* 1(5):191-197, 1987.
291. Hedges, L.V., and Olkin, I., *Statistical Methods for Meta-Analysis* (Orlando, FL: Academic Press, 1985).
292. Heineken, P.A., Charles, G., Stimson, D.H., et al., "The Use of Negative Indexes of Health To Evaluate Quality of Care in a Primary-care Group Practice," *Medical Care* 23(3):198-207, 1985.
293. Heller, T.A., Larson, E.B., and LoGerfo, J.P., "Quality of Ambulatory Care in the Elderly: An Analysis of Five Conditions," *Journal of the American Geriatric Society* 32(11):782-788, 1984.
294. Helmstadter, G.C., *Principles of Psychological Measurement* (New York, NY: Appleton-Century-Crofts, 1964).
295. Hertzer, N.R., Avellone, J.C., Farrell, C.J., et al., "The Risk of Vascular Surgery in a Metropolitan Community," *Journal of Vascular Surgery* 1(1):13-21, 1984.

296. Hibbard, J.H., and Weeks, E.C., "Consumerism in Health Care: Prevalence and Predictors," *Medical Care* 25(11):1019-1032, 1987.

297. Hickson, G.B., Stewart, D.W., Altemeier, W.A., et al., "First Step in Obtaining Child Health Care: Selecting a Physician," *Pediatrics* 81(3):333-338, 1988.

298. Hinshaw, A.S., and Atwood, J.R., "A Patient Satisfaction Instrument: Precision by Replication," *Nursing Research* 31:170-191, 1982.

299. Hinshaw, A.S., Gerber, R.M., Atwood, J.R., et al., "The Use of Predictive Modeling To Test Nursing Practice Outcomes," *Nursing Research* 32:35-42, 1983.

300. Hinz, C.A., "Recredentialing Could Become a Reality for New York Physicians," *American Medical News*, p. 1, July 31, 1987.

301. Hofflander, A.E., and B.F. Nye, *Medical Malpractice Insurance in Pennsylvania* (Menlo Park, CA: Management Analysis Center, Inc., 1985).

302. Holzman, D., "Hospital Death Rates Made Public," *Washington Times*, p. B5, Dec. 22, 1987.

303. Hopkins, K.D., and Stanley, J.C., *Educational and Psychological Measurement and Evaluation* 6th edition (Englewood Cliffs, NJ: Prentice-Hall, 1981).

304. Hsia, D.C., Krushat, W.M., Fagan, A.B., et al., "Accuracy of Diagnostic Coding for Medicare Patients Under the Prospective-Payment System," *New England Journal of Medicine* 318(6):352-355, 1988.

305. Hughes, J.M., Director, Hospital Infections Program, Center for Infectious Diseases, Centers for Disease Control, Public Health Services, U.S. Department of Health and Human Service, Atlanta, GA, personal communication, Nov. 18, 1987.

306. Hughes, R.G., Garnick, D.W., Luft, H.S., et al., "Hospital Volume and Patient Outcomes: The Case of Hip Fracture Patients," *Medical Care*, in press.

307. Hughes, R.G., Hunt, S.S., and Luft, H.S., "Effects of Surgeon Volume and Hospital Volume on Quality of Care in Hospitals," *Medical Care* 25(6):489-503, 1987.

308. Hulka, B.S., Kupper, L.L., and Cassel, J.C., "Physician Management in Primary Care," *American Journal of Public Health* 66(12):1173-1179, 1976.

309. Hulka, B.S., Romm, F.J., Parkerson, G.R., et al., "Peer Review in Ambulatory Care: Use of Explicit Criteria and Implicit Judgments," *Medical Care* 17(3 Suppl.):1-73, 1979.

310. Hulka, B.S., Zyzanski, S.H., Cassel, J.C., et al., "Scale for the Measurement of Attitudes Toward Physicians and Primary Medical Care," *Medical Care* 8:429-436, 1970.

311. Hunter, J.E., Schmidt, F.L., and Jackson, G.B., *Meta-Analysis: Cumulating Research Findings Across Studies* (Beverly Hills, CA: Sage Publishing Co., 1982).

312. Hyman, H.H., "Are Public Hospitals in New York City Inferior to Voluntary, Nonprofit Hospitals? A Study of JCAH Hospital Surveys," *American Journal of Public Health* 76:18-22, 1986.

313. Hypertension Detection and Follow-Up Program Cooperative Group, "Five-Year Findings of the Hypertension Detection and Follow-Up Program," *Journal of the American Medical Association* 242(23):2562-2577, 1979.

314. Inguanzo, J.M., and Harju, M., "What Makes Consumers Select a Hospital?" *Hospitals* 59(4):90-94, 1985.

315. Inguanzo, J.M., and Harju, M., "Affluent Consumers Most Discriminating: Survey," *Hospitals* 59(19):84-86, 1985.

316. Inui, T.S., and Carter, W.B., "Problems and Prospects for Health Services Research on Provider-Patient Communication," *Medical Care* 23:521-538, 1985.

317. Jacobson, P.D., "Medical Malpractice and the Tort System," Rand Corp., Santa Monica, CA, Aug. 18, 1987.

318. Jahnigen, D., Hannon, C., Laxson, L., et al., "Iatrogenic Disease in Hospitalized Elderly Veterans," *Journal of the American Geriatrics Society* 30:387, 1982.

319. Jencks, S.F., and Dobson, A., "Refining Case-Mix Adjustment: The Research Evidence," *New England Journal of Medicine* 317(11):679-686, 1987.

320. Jensen, J., "Choosing a Hospital," *American Demographics* 9(6):45-47, 1987.

321. Jessor, R., "Predicting Time of Onset of Marijuana Use: A Developmental Study of High School Youth," *Predicting Adolescent Drug Abuse: A Review of Issues, Methods and Correlates*, D.J. Lettieri (ed.), National Institute of Drug Abuse, Alcohol and Drug Abuse and Mental Health Administration, Public Health Service, U.S. Department of Health and Human Services (Rockville, MD: 1975).

322. Johns, L., Health Care Planning and Policy, San Francisco, CA, "Quality Counts: Use of a Quality of Care Measure in Development of a PRO," *Business and Health*, in press, July 1988.

323. Joint Commission on the Accreditation of Healthcare Organizations, "Hospital Accreditation Program: Accreditation Decision/Contingency Criteria," Chicago, IL, 1987.

323a. Joint Commission on the Accreditation of Healthcare Organizations, "State Project Status Report," Chicago, IL, Sept. 21, 1987.

324. Joint Commission on the Accreditation of Healthcare Organizations, *National Invitational Forum on Clinical Indicators* (Chicago, IL: November 1987).

325. Joint Commission on the Accreditation of Healthcare Organizations, *AMH/88: Accreditation Manual for Hospitals* (Chicago, IL: 1988).

326. Joint Commission on the Accreditation of Healthcare Organizations, "Monitoring and Evaluating the Quality and Appropriateness of Care," Chicago, IL, Jan. 21, 1988.

327. Joint Commission on the Accreditation of Hospitals, *AMH/86: Accreditation Manual for Hospitals* (Chicago, IL: 1986).

328. Joint Commission on Accreditation of Hospitals, *Hospital Accreditation Program Scoring Guidelines: Medical Staff Standards* (Chicago, IL: 1986).

329. Joint Commission on the Accreditation of Hospitals, "The Joint Commission's 'Agenda for Change,' " Chicago, IL, November 1986.

330. Joint Commission on the Accreditation of Hospitals, *AMA/87: Accreditation Manual for Hospitals 1987* (Chicago, IL: 1987).

331. Joint Commission on the Accreditation of Hospitals, "Hospital Accreditation Program: Accreditation Decision/Contingency Criteria," Chicago, IL, 1987.

332. Joint National Committee on Detection, Evaluation, and Treatment of High Blood Pressure, "Hypertension Prevalence and the Status of Awareness, Treatment, and Control in the United States," *Hypertension* 7(3):457-468, 1985.

333. Jost, T.S., "The Joint Commission on Accreditation of Hospitals: Private Regulation of Health Care and the Public Interest," *Boston College Law Review* 24(4):853-857, July 1983.

334. Journal of the American Medical Association, "National Institutes of Health Consensus Development Conference: Critical Care Medicine," *Journal of the American Medical Association* 250:798, 1983.

335. Journal of the American Medical Association, "AMA Initiative on Quality of Medical Care and Professional Self-Regulation," *Journal of the American Medical Association* 256(8):1036-1037, 1986.

336. Journal of the American Medical Association, "Certification: Another Step for Geriatric Medicine," *Journal of the American Medical Association* 258(11):1518, 1987.

337. Kahneman, D., and Tversky, A., "The Psychology of Preferences," *Scientific American* 246(1):160-173, 1982.

338. Kalven, H., and Zeisel, H., *The American Jury* (Boston, MA: Little Brown, 1966).

338a. Kandel, D., "Some Comments on the Relationship of Selected Criterion Variables to Adolescent Drug Use," *Predicting Adolescent Drug Abuse: A Review of Issues, Methods and Correlates*, D.J. Lettieri (ed.), National Institute of Drug Abuse, Alcohol and Drug Abuse and Mental Health Administration, Public Health Service, U.S. Department of Health and Human Services (Rockville, MD: 1975).

339. Kane, R.L., Gardner, J., Wright, D.D., et al., "Relationship Between Process and Outcome in Ambulatory Care," *Medical Care* 15(11):961-965, 1977.

340. Kane, R.L., Klein, S.J., Bernstein, L., et al., "Hospice Role in Alleviating the Emotional Stress of Terminal Patients and their Families," *Medical Care* 23:189-197, 1985.

341. Kasteler, J., Kane, R.L., Olsen, D.M., et al., "Issues Underlying Prevalence of 'Doctor-Shopping' Behavior," *Journal of Health and Social Behavior* 17(4):329-339, 1976.

342. Katz, D., "What the State Revokes, the Courts Can Restore," *The Detroit Free Press*, p.13A, Apr. 1, 1984.

343. Keeler, E.B., Brook, R.H., Goldberg, G.A., et al., *How Free Care Reduced Hypertension of Participants in the Rand Health Insurance Experiment* (Santa Monica, CA: Rand Publication Series, 1985).

344. Keeswater, B., Chief of Acute Care Unit, Acute Care Hospital Regulation, State of California Department of Health Services, Sacramento, CA, personal communication, Jan. 19, 1988.

345. Keller, P., Chief of Field Operations Branch, Licensing and Certification, State Department of Health Services, Sacramento, CA, personal communication, Jan. 8, 1988.

346. Kelly, J.V., "Deaths in Teaching and Nonteaching Hospitals: Quality of Care and Severity of Illness Differences," forthcoming.

347. Kelly, J.V., and Hellinger, F.J., "Physician and Hospital Factors Associated With Mortality of

Surgical Patients," *Medical Care* 24(9):785-800, 1986.

348. Kelly, J.V., and Hellinger, F.J., "Heart Disease and Hospital Deaths: An Empirical Study," *Health Services Research* 22(3):369-395, 1987.

349. Kempczinski, R.F., Brott, T.G., and Labutta, R.J., "The Influence of Surgical Specialty and Caseload on the Results of Carotid Endarterectomy," *Journal of Vascular Surgery* 3(6):911-916, 1986.

350. Kempthorne, G., Medical Director, HMO of Wisconsin, Spring Green, WI, personal communication, February 1988.

351. Kessner, D.M., Kalk, C.E., and Singer, I., "Assessing Health Quality—The Case for Tracers," *New England Journal of Medicine* 288(4):189-194, 1973.

352. Knaus, W.A., Draper, E.Q., Wagner, D.P., et al., "APACHE II—A Severity of Disease Classification System," *Critical Care Medicine* 13:818-829, 1985.

353. Knaus, W.A., Draper, E.Q., Wagner, D.P., et al., "An Evaluation of Outcome From Intensive Care in Major Medical Centers," *Annals of Internal Medicine* 104:410-418, 1986.

354. Knudson, M., "Death Rate Found to Vary for Digestive Operations," *Baltimore Sun*, p. 1A, Nov. 22, 1987.

355. Komaroff, A.L., "Quality Assurance in 1984," *Medical Care* 23(5):723-734, 1985.

356. Kowalski, G., Cancer Department, American College of Surgeons, Chicago, IL, personal communication, Mar. 25, 1988.

357. Krakauer, H., Office of Medical Review, Health Standards and Quality Bureau, Health Care Financing Administration, U.S. Department of Health and Human Services, Baltimore, MD, personal communication, Feb. 1, 1988, Feb. 19, 1988, and Mar. 7, 1988.

358. Krames Communications, "Asking Questions: For Only the *Best* Health Care," Daly City, CA, 1986.

359. Krieger, L.M., and Rosenfeld, S., "Study Compares Death Rates at Hospitals," *San Francisco Examiner*, p. 1, Apr. 24, 1987.

360. Kusserow, R.P., Inspector General, U.S. Department of Health and Human Services, testimony at hearing on the Medicare Quality Assurance Process before the Subcommittee on Intergovernmental Relations and Human Resources, Committee on Government Operations, House of Representatives, U.S. Congress, Washington, DC, Oct. 20, 1987.

361. Kusserow, R.P., Handley, E.A., and Yessian, M.R., "An Overview of State Medical Discipline," *Journal of the American Medical Association* 257(6):821-824, 1987.

362. Lambert, B., "New York Panel Urges Re-Testing of All Physicians," *New York Times*, p. A1 (col. 1), Feb. 26, 1988.

363. LaMonica, E.L., Oberst, M.T., Madea, A.R., et al., "Development of a Patient Satifaction Scale," *Research in Nursing and Health* 9(1):43-50, 1986.

364. Langsley, D.G., "Prior ABMS Conferences on Recertification," *Recertification for Medical Specialists*, J.S. Lloyd and D.G. Langsley (eds.) (Evanston, IL: American Board of Medical Specialties, 1987).

365. Langsley, D.G., Executive Vice President, American Board of Medical Specialties, Evanston, IL, Feb. 5, 1988.

366. Larks, M., "Access to Health Data by State Health Data Organizations and Quality Assessors," contractor document prepared for the Office of Technology Assessment, U.S. Congress, Washington, DC, December 1987.

367. LaTour, S.A., Marketing Department, J.L. Kellogg Graduate School of Management, Northwestern University, Evanston, IL, personal communication, May 14, 1987.

368. Lebow, J., "Consumer Assessments of the Quality of Medical Care," *Medical Care* 12:328-337, 1974.

369. Lefebre, M., Zweig, S., and Druse, J., "Selection of a Physician for Prenatal Care," *Journal of Family Practice* 24(3):275-279, 1987.

370. LeMaitre, G.D., *How To Choose a Good Doctor* (Andover, MA: Andover Publishing Group, 1979).

371. Lembcke, P.A., Evaluation of Medical Audit," *Journal of the American Medical Association* 199:543-550, Feb. 20, 1967.

372. Lenfant, C., "Advancements in Meeting the 1990 Hypertension Objectives," *Morbidity and Mortality Weekly Report* 36(10):144-151, 1987.

373. Levitt, T., "Marketing Intangible Products and Product Intangibles," *Harvard Business Review* 59(3):94-102, 1981.

374. Ley, P., Kinsey, J., and Atherton, S.T., "Increasing Patients' Satisfaction With Communication," *British Journal of Social and Clinical Psychology* 15:403-413, 1976.

375. Libercci, W., Deputy Director of Health Care Administrative Sanctions Staff, Office of the Inspector General, Office of Investigations, Health Care Financing Administration, U.S. Department of Health and Human Services, Baltimore, MD, personal communication, Oct. 28, 1987.

376. Light, R., and Pillemer, D., *Summing Up: The Science of Reviewing Research* (Cambridge, MA: Harvard University Press, 1984).

377. Linn, L.S., "Factors Associated With Patient Evaluation of Health Care," *Milbank Memorial Fund Quarterly* 53:531-548, 1975.

378. Linn, L.S., "Burn Patients' Evaluation of Emergency Department Care," *Annals of Emergency Medicine* 11:255-259, 1982.

379. Linn, L.S., and Greenfield, S., "Patient Suffering and Patient Satisfaction Among the Chronically Ill," *Medical Care* 20:425-431, 1982.

380. Linton, O., Executive Director, American College of Radiology, Bethesda, MD, Dec. 8, 1987.

381. Loft, J.D., and Ryan, G.A., "The American Medical Association's Physician Masterfile," *The 1985 Public Health Conference on Records and Statistics*, Aug. 13-15, 1985 (Hyattsville, MD: U.S. Department of Health and Human Services, Public Health Service, National Center for Health Statistics, 1985).

382. Lohr, K.N., "Peer Review Organizations (PROs): Quality Assurance in Medicare," prepared for the Office of Technology Assessment, U.S. Congress, Washington, DC, November 1984.

383. Lohr, K.N., "Commentary: Professional Peer Review in a 'Competitive' Medical Market," *Case Western Review Law Review* 36(4):1175-1189, 1985-1986.

384. Lohr, K.N., and Brook, R.H., "Quality Assurance in Medicine," *American Behavioral Scientist* 27(5):583-607, 1984.

385. Lohr, K.N., Brook, R.H., and Kaufman, M.A., "Quality of Care in the New Mexico Medicaid Program (1971-1975)," *Medical Care* 18(1):1-103, 1980.

386. Lohr, K.N., Brook, R.H., Kamberg, C.J., et al. "Use of Medical Care in the Rand Health Insurance Experiment: Diagnosis- and Service-Specific Analyses in a Randomized Controlled Trial," *Medical Care* 24(9 Suppl.):S1-87, 1986.

387. Longo, D.R., Roberts, J.S., and Wilt, J.E., "Compliance With JCAH Standards: A Comparison of 1982 and 1983 Surveys," *Quality Review Bulletin* 11(12):385-388, 1985.

388. Longo, D.R., Wilt, J.E., and Laubenthals, R.M., "Hospital Compliance With Joint Commission Standards: Findings From 1984 Surveys," *Quality Review Bulletin* 12(11):388-393, 1986.

389. Louie, S., Legal Counsel, Board of Medical Quality Assurance, Sacramento, CA, personal communication, Sept. 21, 1987.

390. Louis Harris and Associates, *Hospital Care in America* (New York, NY: April 1978).

391. Louis Harris and Associates, "What New Yorkers Say About Their Health Care," New York, NY, 1985.

392. Louis Harris and Associates, *Americans and Their Doctors* (New York, NY: January 1985).

393. Luft, H.S., "The Relation Between Surgical Volume and Mortality: An Exploration of Causal Factors and Alternative Models," *Medical Care* 18(9):940-959, 1980.

394. Luft, H.S., Bunker, J.P., and Enthoven, A.C., "Should Operations Be Regionalized: The Empirical Relation Between Surgical Volume and Mortality," *New England Journal of Medicine* 301(25):1364-1369, Dec. 20, 1979.

395. Luft, H.S., Garnick, D.W., Mark, D., et al., "Evaluating Research on the Use of Volume of Services Performed in Hospitals as an Indicator of Quality," contractor document prepared for the Office of Technology Assessment, U.S. Congress, Washington, DC, December 1987.

396. Luft, H.S., and Hunt, S.S., "Evaluating Individual Hospital Quality Through Outcome Statistics," *Journal of the American Medical Association* 255(20):2780-2784, 1986.

397. Luft, H.S., Hunt, S.S., and Maerki, S.C., "The Volume-Outcome Relationship: Practice Makes Perfect or Selective Referral Patterns?" *Health Services Research* 22(2):157-182, 1987.

398. Lurie, N., Ward, N.B., Shapiro, M.F., et al., "Termination From Medi-Cal: Does it Affect Health?" *New England Journal of Medicine* 311:480-484, 1984.

399. Luten, E., Surveyor Training Branch, Health Care Financing Administration, U.S. Department of Health and Human Services, Baltimore, MD, "Improve Surveyor Training," Sept. 11, 1987.

400. Lyons, T.F., and Payne, B.C., "The Relationship of Physicians' Medical Recording Performance to their Medical Care Performance," *Medical Care* 12(5):463-469, 1974.

401. Madnick, M.E., *Consumer Health Education: A Guide to Hospital-Based Programs* (Wakefield, MA: Nursing Resources, 1980).

402. Maerki, S.C., Luft, H.S., and Hunt, S.S., "Selecting Categories of Patients for Regionalization: Implications of the Relationship Between Volume and Outcome," *Medical Care* 24(2):148-158, 1986.

403. Mankin, H.J., Leffert, R.D., Schachar, N.S., et al., "The Use of the Orthopaedic Educational

Computer Databank (OECD) To Analyze the Content of the Orthopaedic In-Training Examinations," *Evaluating the Skills of Medical Specialists* (Chicago, IL: American Board of Medical Specialties, 1983).

404. Marketing News, "Psychographics Help Health Marketers Find and Serve New Market Segments," *Marketing News*, p. 4, Apr. 24, 1987.

405. Maroc, J., M.D., Director of Medical Affairs, Iowa Foundation for Medical Care, Peer Review Organization, West Des Moines, IA, personal communication, Oct. 16, 1987.

406. Marquis, M.S., Davies, A.R., and Ware, J.E., Jr., "Patient Satisfaction and Change in Medical Care Provider: A Longitudinal Study," *Medical Care* 21:821-829, 1983.

407. Marquis, M.S., Kanouse, D.E., and Brodsley, L., *Informing Consumers About Health Care Costs: A Review and Research Agenda*, prepared for the Health Care Financing Administration, U.S. Department of Health and Human Services (Santa Monica, CA: Rand Corp., September 1985).

408. Maryland Hospital Association, "Hospital Mortality Data Analysis," Luterville, MD, 1987.

409. Massanari, R.M., Wilkerson, K., Streed, S.A., et al., "Reliability of Reporting Nosocomial Infections in the Discharge Abstract and Implications for Receipt of Revenues Under Prospective Reimbursement," *American Journal of Public Health* 77(5):561-564, 1987.

410. McAuley, R.G., and Henderson, H.W., "Results of the Peer Assessment Program of the College of Physicians and Surgeons of Ontario," *Canadian Medical Association Journal* 131:557-561, 1984.

411. McAuliffe, W., "Studies of Process-Outcome Correlations in Medical Care Evaluations," *Medical Care* 16(11):907-930, 1978.

412. McAuliffe, W., "A Validation Theory for Quality Assessment," in J.J. Pena et al., *Risk Management and Program Evaluation: Tools for Clinicians and Administrators* (Aspen, CO: Aspen Press, 1983).

413. McAuliffe, W., Department of Psychiatry, Cambridge Hospital, Cambridge, MA, personal communication, Apr. 15, 1987.

414. McBarnette, L., Deputy Director, New York State Department of Health, Albany, NY, personal communication, 1987.

415. McClellan, M., "Hypertension Screening and Management as an Indicator of Quality: An Evaluation of the Literature," working paper, Office of Technology Assessment, U.S. Congress, Washington, DC, March 1988.

416. McCosker, M., Private HealthCare Systems, Lexington, MA, personal communication, Oct. 9, 1987.

417. McCoy, C.E., Kind, E.A., Fowles, J., et al., "Measuring Quality in an HMO: The Primary Care Practice Profile," Park Nicollet Medical Foundation, Minneapolis, MN, February 1988.

418. McCoy, M., School of Medicine, University of California at Los Angeles, Los Angeles, CA, personal communication, July 1987.

419. McCoy, M., School of Medicine, University of California at Los Angeles, presentation at the American Public Health Association Meeting, New Orleans, LA, Oct. 18, 1987.

420. McDonough, W.J., Commonwealth of Massachusetts Board of Registration in Medicine, Boston, MA, personal communication, Dec. 3, 1987.

421. McGhan, W.F., Stimmel, G.L., Hall T.G., et al., "A Comparison of Pharmacists and Physicians on the Quality of Prescribing for Ambulatory Hypertensive Patients," *Medical Care* 21(4):435-444, 1983.

422. McGuire, W.J., "The Current Status of Cognitive Consistency Theories," *Behavioral Science Foundations of Consumer Behavior*, J.B. Cohen (ed.) (New York, NY: The Free Press, 1972).

423. Medical World News, "AMA To Weed Out Its Bad Apples," *Medical World News* 27(14):30, 1986.

424. Medical World News, "Hospitals Not in Step on Infection Control," *Medical World News* 27(22):7, 1986.

425. Medical World News, "Hypertension Experts See More Stepped-Care Treatment Choices," *Medical World News* 28(11):14-15, 1987.

426. Medical World News, "Hospitals Brace for Death Rate Releases," *Medical World News* 28(15):10-19, 1987.

427. Medical World News, "Pennsylvania Plans To Release Doctor and Hospital Outcome Data," *Medical World News* 28(23):8-9, 1987.

428. Medicine and Health, "Peer Review of Ambulatory Care," *Medicine and Health* 41(35):insert, 1987.

429. Mendelsohn, P., Branch Chief, Project Assessment Branch, Office of Medical Review, Health Care Financing Administration, U.S. Department of Health and Human Services, Baltimore, MD, Apr. 7, 1988.

430. Meyers, A.R., " 'Lumping It': The Hidden Denominator of the Medical Malpractice Crisis," *American Journal of Public Health* 77(12):1544-1548, 1987.

431. Miller, G.A., "The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information," *Psychological Review* 63:81-97, March 1956.

432. Mills, D.H. (ed.), *Report on the Medical Insurance Feasibility Study* (San Francisco, CA: California Medical Association, 1977).

433. Mills, D.H., and Lindgren, O., "Physician Malpractice Profiles as Indicators of Quality: Reliability, Validity, and Feasibility Issues," contractor document prepared for the Office of Technology Assessment, U.S. Congress, Washington, DC, November 1987.

434. Mitchner, B., "Physician Specialization as an Indicator of Quality: An Evaluation of the Literature," working paper, Office of Technology Assessment, U.S. Congress, Washington, DC, 1988.

435. Moncrief, W.H., President, California Medical Review, Inc., San Francisco, CA, personal communication, Feb. 19, 1987, and Apr. 18, 1988.

436. Moncrief, W.H., California Medical Review, Inc., testimony at hearing on Medicare's Peer Review Organizations before the Subcommittee on Intergovernmental Relations and Human Resources, Committee on Government Operations, House of Representatives, U.S. Congress, Oct. 20, 1987.

437. Monroe, K.B., and Krishnan, R., "The Effect of Price on Subjective Product Evaluation," *Consumers of Merchandise and Store Quality*, J. Jacoby and J. Olson (eds.) (Lexington, MA: D.C. Heath, 1984).

438. Moran, M., Director, Division of Data and Program Analysis, Health Standards and Quality Bureau, Health Care Financing Administration, U.S. Department of Health and Human Services, Baltimore, MD, personal communication, Sept. 24, 1987, and Jan. 20, 1988.

439. Morbidity ~~and~~ Mortality Weekly Report, "Premature Mortality in the United States," *Morbidity and Mortality Weekly Report* 35(2S):1S-11S, Dec. 19, 1986.

440. Morbidity and Mortality Weekly Report, "Premature Mortality Due to Unintentional Injuries—United States, 1984," *Morbidity and Mortality Weekly Report* 36(49):814-815, 1987.

441. Moses, L., and Mosteller, F., "Institutional Differences in Postoperative Death Rates: Commentary on Some of the Findings of the National Halothane Study," *Journal of the American Medical Association* 203:492, 1968.

442. Mosteller, F., "Assessing Quality of Institutional Care," *American Journal of Public Health* 77(9):1155-1156, 1987.

443. Mosteller, F., Chairman, Department of Health Policy and Management, Harvard School of Public Health, Boston, MA, personal communication, December 1987.

444. Moynihan, C., and Barnes, C., "Accuracy of HCFA Generic Screens in Identifying Quality Problems: Analysis of False Positive and False Negative Occurrences," paper presented at the American Medical Review Research Center Second Annual Research Symposium, Washington, DC, Sept. 10, 1987.

445. Mullen, P.E., Director, Government Relations, Joint Commission on Accreditation of Healthcare Organizations, Chicago, IL, personal communication, 1988.

446. Munier, W., Quality Standards in Medicine, Inc., Wellesley, MA, presentation at the American Public Health Association Meeting, New Orleans, LA, Oct. 18, 1987.

447. Munier, W., Quality Standards in Medicine, Inc., Wellesley, MA, personal communication, Dec. 3, 1987.

448. National Academy of Sciences, *Health Care for American Veterans*, Committee on Veterans' Affairs Print No. 36, prepared for House of Representatives, U.S. Congress, 95th Congress, 1st Session (Washington, DC: U.S. Government Printing Office, 1977).

449. National Academy of Sciences, Institute of Medicine, *Beyond Malpractice: Compensation for Medical Injuries* (Washington, DC: March 1978).

450. National Academy of Sciences, Institute of Medicine, "Reliability of National Hospital Discharge Survey Data," Washington, DC, 1980.

451. National Academy of Sciences, Institute of Medicine, *Access to Medical Review Data* (Washington, DC: National Academy Press, October 1981).

452. National Academy of Sciences, Institute of Medicine, *Planning Study Report: A Consortium for Assessing Medical Technology* (Washington, DC: National Academy Press, 1983).

453. National Academy of Sciences, Institute of Medicine, *Assessing Medical Technologies* (Washington, DC: National Academy Press, 1985).

454. National Academy of Sciences, National Research Council, *Accidental Death and Disability: The Neglected Disease of Modern Society* (Washington, DC: Government Printing Office, 1966).

455. National Academy of Sciences, National Research Council and Institute of Medicine, *Injury in America* (Washington, DC: National Academy Press, 1985).

456. National Board of Medical Examiners, *Annual Report* (Philadelphia, PA: 1987).

457. National Emergency Medical System Clearinghouse, *State Trauma Center Programs: the Current Status* (Lexington, KY: July 1986).

458. National Women's Health Network, "Plaintext Doctor-Patient Checklist," Washington, DC, circa 1985.

459. Neuhaus, E., Lyons, T.F., and Payne, B.C., "Problems of Case Finding and Data Collection in Ambulatory Care Settings," *American Journal of Public Health* 70(3):282-283, 1980.

460. New York State Department of Health, *Monitoring Health Care Quality, Malpractice Misconduct Quality Assurance* (Albany, NY: March 1988).

461. New York State Department of Health, Office of Health Systems Management, "Investigation of Quality of Care Problems Associated With In-Hospital Mortality in New York State: An Identification of Critical Case Characteristics for Targeting Purposes," Albany, NY, December 1987.

462. New York State Department of Health, Office of Health Systems Management, "Investigation of Quality of Care in Hospitals," Albany, NY, 1987.

463. Newhouse, J.P., Manning, W.G., Morris, C.N., et al., "Some Interim Results From a Controlled Trial of Cost Sharing in Health Insurance," *New England Journal of Medicine* 305(25):1501-1507, 1981.

464. Newhouse, J.P., Ware, J.E., Jr., and Donald, C.A., "How Sophisticated are Consumers About the Medical Care Delivery System?" *Medical Care* 19(3):316-328, 1981.

465. Nobrega, F.T., Morrow, G.W., Smoldt, R.K., et al., "Quality Assessment in Hypertension: Analysis of Process and Outcome Methods," *New England Journal of Medicine* 296(3):145-148, 1977.

466. Norcini, J.J., Swanson, D.B., and Webster, G.D., "Reliability, Validity and Efficiency of Various Item Formats in Assessment of Physician Competence," *The Annual Conference on Research in Medical Education, 1983* (Washington, DC: 1983).

467. Nunnally, J.C., *Psychometric Theory*, 2nd edition (New York, NY: McGraw Hill Book Co., 1978).

468. Nutting, P.A., Burkhalter, B.R., Dietrick, D.L., et al., "Relationship of Size and Payment Mechanism to System Performance in 11 Medical Care Systems," *Medical Care* 20(7):676-690, 1982.

469. Nye, B.F. (President of Stanford Consulting Group, Inc.), and Hofflander, A.E. (Professor of Finance and Insurance at the Graduate School of Management, University of California, Los Angeles), "Experience Rating in Medical Professional Liability Insurance," 1986.

469a. Opaluch, G.M., Approvals Program Administrator, American College of Surgeons, Commission on Cancer, Chicago, IL, personal communication, Mar. 30, 1988.

470. Osteen, A., "Medical Licensing Requirements," *Journal of the American Medical Association* 258(8):1053-1054, 1987.

471. Osteen, A., American Medical Association, Chicago, IL, personal communication, Feb. 19, 1987.

472. Osten, W.M., Director, Bureau of Hospital Services, Office of Health Systems Management, State of New York Department of Health, Albany, NY, personal communication, December 1987.

473. Page, L., "Specialists Debate: Who Should Do Liposuction," *American Medical News*, p. 1, Aug. 21, 1987.

474. Palmer, D.B., "Some Cons of PROs," *Archives of Ophthalmology* 103:343-345, 1985.

475. Palmer, R.H., *Ambulatory Health Care Evaluation: Principles and Practice* (Chicago, IL: American Hospital Publishing, Inc., 1983).

476. Palmer, R.H., Lecturer in Health Services, Harvard School of Public Health and Institute for Health Research, Boston, MA, personal communication, September 1987.

477. Palmer, R.H., and Reilly, M.C., "Individual and Institutional Variables Which May Serve as Indicators of Quality of Medical Care," *Medical Care* 17(7):693-717, 1979.

478. Patten, B., "Open Market, Open Heart: Special Report," *Phoenix [Arizona] Gazette*, p. A1, Aug. 26, 1987.

479. Payne, B.C., "The Medical Record as a Basis for Assessing Physician Competence," *Annals of Internal Medicine* 91(4):623-629, 1979.

480. Payne, B.C., and Lyons, T.F., *Office Care Study* (Ann Arbor, MI: University of Michigan School of Medicine, 1972).

481. Payne, B.C., Lyons, T.F., and Neuhaus, E., "Relationships of Physician Characteristics to Performance Quality and Improvement," *Health Services Research* 19(3):307-332, 1984.

482. Pekannen, J., "Good Family Doctors," *Washingtonian* 23(1):115-120, 1987.

483. Pekannen, J., "How To Choose a Doctor: Advice From the Experts on Making the Right Choice," *Washingtonian* 23(1):122, October 1987.

484. Pennsylvania Health Care Cost Containment Council, "Data Elements as Approved by the Pennsylvania Health Care Cost Containment Council," Harrisburg, PA, May 7, 1987.

485. Pennsylvania Health Care Cost Containment Council, Harrisburg, PA, proposals for "Data Reporting Requirements" and "Quality Data Reporting Requirements" submitted to Pennsylvania's Independent Regulatory Review Commission, Sept. 7, 1987, revised Nov. 8, 1987.

486. Peterson, M.A., *Civil Juries in the 1980's: Trends in Jury Trials and Verdicts in California and Cook County, Illinois*, R03466-ICJ (Santa Monica, CA: Rand Corp., 1987).

487. Philipsen, M.A., Office of Medical Review, Health Standards and Quality Bureau, Health Care Financing Administration, U.S. Department of Health and Human Services, Baltimore, MD, personal communication, Nov. 30, 1987.

488. Pilcher, D.B., Davis, J.H., Ashikaga, T., et al., "Treatment of Abdominal Aortic Aneurysm in an Entire State Over 7 1/2 Years," *American Journal of Surgery* 139(4):487-494, 1980.

489. Pillemer, D. and Light, R., "Synthesizing Outcomes: How To Use Research Evidence From Many Studies," *Harvard Educational Review* 50:176-195, 1980.

490. Pisacano, N.J., Executive Director and Secretary, American Board of Family Practice, Lexington, KY, personal communication, May 8, 1987.

491. Pitman, L.W., Chief Executive Officer, Kansas Foundation for Medical Care, testimony at hearing on Medical Peer Review Organization Program before the Subcommittee on Health and the Environment, Committee on Energy and Commerce, House of Representatives, U.S. Congress, Washington, DC, Oct. 26, 1987.

492. Platt, K.A., Medical Director, Colorado Foundation for Medical Care, "Peer Review Organizations and Quality Assurance," presentation at National Academy of Sciences, Institute of Medicine, Council on Health Care Technology Forum Program: Quality of Care and Technology Assessment, Washington, DC, May 14-16, 1987.

493. Platt, R., Polk, B.F., Murdock, B., et al.,"Mortality Associated With Nosocomial Urinary-Tract Infection," *New England Journal of Medicine* 307(11):637-642, 1982.

494. Pocincki, L.S., Dogger S.J., Schwartz, B.P., "The Incidence of Iatrogenic Injuries," Appendix to *Report of the Secretary's Commission on Medical Malpractice*, DHEW Pub. No. [OS]:73-89 (Washington, DC: U.S. Department of Health, Education, and Welfare, 1973).

495. Pollner, F., "The Angioplasty Debate Grows With Its Use," *Medical World News* 28(1):52-72, 1987.

496. Powills, S., "Who's Making Health Care Decisions?" *Hospitals* 61(15):40, 1987.

497. Prevost, J.A., Director for Research and Development, Joint Commission on the Accreditation of Healthcare Organizations, remarks made at National Invitational Forum on Clinical Indicators, Chicago, IL, November 1987.

498. Project HOPE, Center for Health Affairs, "Buying Quality Health Care: What to Ask When Purchasing Health Insurance and Medical Services," Chevy Chase, MD, The People-to-People Health Foundation, Inc., 1987.

499. Public Citizen Health Research Group, "State Medical Licensing Board Disciplinary Actions in 1985: Better, But Not Good Enough," Washington, DC, Nov. 10, 1986.

500. Public Citizen Health Research Group as cited in *Medicine and Health*, "Consumer Group Urges HCFA To Improve Hospital Mortality Ratings," Sept. 14, 1987.

501. Public Citizen Health Research Group, "Medicare Update," *Public Citizen Health Research Group Health Letter*, Washington, DC, October 1987.

502. Public Citizen Health Research Group, "Press Guide to HCFA Mortality Data Release," Washington, DC, Dec. 10, 1987.

503. Public Citizen Health Research Group, "Hospital Mortality Rates," *Public Citizen Health Research Group Health Letter* 4(2):1-8, Washington, DC, 1988.

504. *Purcell v. Zimbelman*, 500 P.2d 335 (Ct. App. Ariz., 1972)
505. Quebec, *Report of the Commission of Inquiry on Health and Social Welfare, Tome III, Vol. IV, Part 2, Second Title: The Health Plan* (Quebec City: Government of Quebec, 1970).
506. Quinn, S., Staff Specialist, Resource Center, American Hospital Association, Chicago, IL, personal communication, March 1988.
507. Rakel, R.E., "Defining Competence in Specialty Practice: The Need for Relevance," *Evaluating the Skills of Medical Specialists* (Chicago, IL: American Board of Medical Specialties, 1983).
508. Ramenofsky, M.L., and Morse, T.S., "Standards of Care for the Critically Injured Pediatric Patient," *Journal of Trauma* 22:921, 1982.
509. Ramsay, J.A., McKenzie, J.K., and Fish, D.G., "Physicians and Nurse Practitioners: Do They Provide Equivalent Health Care?" *American Journal of Public Health* 72(1):55-57, 1982.
510. Ramsey, P.G., Carline, J.D., Inui, T.S., et al., "Assessment of the Clinical Competence of Practicing Internists," final report to the American Board of Internal Medicine, Philadelphia, PA, 1986.
511. Reade, J.M., and Ratzan, R.M., "Yellow Professionalism," *New England Journal of Medicine* 316(21):1315-1319, 1987.
512. Rees, A.M., and Young, B.A., *The Consumer Health Information Source Book* (New York, NY: R.R. Bowker Co., 1981).
513. Resident & Staff Physician, "How To Get Your License: A State-by-State Guide," *Resident & Staff Physician* 32(5):89-124, 1986.
514. Rhee, S., "Factors Determining the Quality of Physician Performance in Patient Care," *Medical Care* 14(9):733-749, 1976.
515. Rhee, S., "Relative Importance of Physicians' Personal and Situational Characteristics for the Quality of Patient Care," *Journal of Health and Social Behavior* 18(3):10-15, 1977.
516. Rhee, S., Roice, D.L., Lyons, T.F., et al., "Domain of Practice and the Quality of Physician Performance," *Medical Care* 19(1):14-23, 1981.
517. Rice, C.E., Berger, D.G., Klett, S.L., et al., "The Ward Evaluation Scale," *Journal of Clininical Psychology* 19:251-258, 1963.
518. Richardson, F.M., "Methodological Development of a System of Medical Audit," *Medical Care* 10:451-462. 1972.
519. Riley, G., and Eggers, P., "Evaluation of the Impact of Prospective Payment on Quality of Care," paper presented at the American Public Health Association Meeting, Las Vegas, NV, Sept. 29, 1986.
520. Riley, G., and Lubitz, J., "Outcomes of Surgery Among the Medicare Aged: Surgical Volume and Mortality," *Health Care Financing Review* 7(1):37-47, 1985.
521. Rivers, S., Director, Division of State Legislation, American Medical Association, Chicago, IL, personal communication, Sept. 22, 1987.
522. Robert Wood Johnson Foundation, "First Round of Grants Announced Under Medical Malpractice Program," news release, Princeton, NJ, June 29, 1987.
523. Roberts, J.S., "Research Questions," prepared for the Steering Committee on the Survey and Accreditation Project, Joint Commission on the Accreditation of Hospitals, Chicago, IL, December 1986.
524. Roberts, J.S., Senior Vice President for Research and Planning, Joint Commission on the Accreditation of Healthcare Organizations, Chicago, IL, personal communication, 1987, Feb. 19, 1987, and Feb. 16, 1988.
525. Roemer, M.R., and Friedman, J.W., *Doctors in Hospitals: Medical Staff Organization and Hospital Performance* (Baltimore, MD: Johns Hopkins University Press, 1971).
526. Roemer, M.R., Moustafa, A.T., and Hopkins, C.E., "A Proposed Hospital Quality Index: Hospital Death Rates Adjusted for Case Severity," *Health Services Research* 3:96, 1968.
527. Rogers, J.L., Haring, O.M., Wortman, P.M., et al., "Medical Information Systems: Assessing Impact in the Areas of Hypertension, Obesity, and Renal Disease," *Medical Care* 20(1):63-74, 1982.
528. Rojek, D.G., Clemente, F., and Summers, C.F., "Community Satisfaction: A Study of Contentment With Local Services," *Rural Sociology* 40:177-192, 1975.
529. Rolph, J.E., "Some Statisical Evidence On Merit Rating in Medical Malpractice Insurance," *Journal of Risk and Insurance* 68(2):247-260, 1981.
530. Romm, F.J., and Hulka, B.S., "Peer Review in Diabetes and Hypertension: The Relationship Between Care Process and Patient Outcome," *Southern Medical Journal* 73(5):564-568, 1980.
531. Roos, L.L., Cageorge, S.M., Roos, N.P., et al., "Centralization, Certification, and Monitoring Readmissions and Complications After Surgery," *Medical Care* 24(11):1044-1066, 1986.
532. Roos, L.L., Roos, N.P., Cageorge, S.M., et al., "How Good Are the Data? Reliability of One Health Care Data Bank," *Medical Care* 20(3):266-276, 1982.
533. Roos, L.L., Roos, N.P., and Sharp, S.M., "Monitoring Adverse Outcomes of Surgery

Using Administrative Data," *Health Care Financing Review*, Annual Supplement, pp. 5-16, December 1987.

534. Roper, W.L., "Legislative Update: Quality Mechanisms," *Health Care Financing Review* 8(1):87-91, Fall 1986.

535. Roper, W.L., Administrator, Health Care Financing Administration, testimony at hearing on Medicare's Peer Review Organizations before the Subcommittee on Health and the Environment, Committee on Energy and Commerce, House of Representatives, U.S. Congress, Oct. 26, 1987.

536. Rose, R., Hunting, K.J., Townsend, T.R., et al., "Morbidity/Mortality and Economics of Hospital-Acquired Blood Stream Infections: A Controlled Study," *Southern Medical Journal* 70:1267-1269, 1977.

537. Rosen, H., and Green, B.A., "The HCFA Excess Mortality Lists: A Methodological Critique," *Hospital & Health Services Administration* 32(1):119-127, February 1987.

538. Rosenblatt, R.A., Reinken, J., and Shoemack, P., "Is Obstetrics Safe in Small Hospitals? Evidence From New Zealand's Regionalised Perinatal System," *Lancet* 2(8452):429-432, 1985.

539. Rosenthal, R., *Meta-Analytic Procedures for Social Research* (Beverly Hills, CA: Sage Publishing Co., 1984).

540. Roter, D.L., "Patient Participation in the Patient-Provider Interaction: The Effects of Patient Question Asking on the Quality of Interaction, Satisfaction, and Compliance," *Health Education Monographs* 5(4):281-315, 1977.

541. Rothschild, M.L., "Advertising Strategies for High and Low Involvement Situations," *Attitude Research Plays for High Stakes*, J.C. Maloney and B. Silverman (eds.) (Chicago, IL: American Marketing Association, 1979).

542. Rovner, S., "The Elderly, Flexing their Medical Muscle," *Washington Post*, Health Section, p. 8, Mar. 12, 1986.

543. Roy, P.D., "The Value of Trauma Centres: A Methodologic Review," *Canadian Journal of Surgery* 30:17, 1987.

544. Rubin, H.R., Clinical Scholar, University of California at Los Angeles, "Using Information From Patients in Evaluating the Quality of Hospital Care: A Literature Review," forthcoming.

545. Rust, F.P., Rust, K.J., and Williams, R.L., Health Data Research Facility, Community and Organization Research Institute, University of California, *1980-1984 Maternal and Child Health Data Base: Descriptive Narrative*, Contract No. MCH 86-89089, report for the Maternal and Child Health Branch, Department of Health Services, State of California (Santa Barbara, CA: Regents of the University of California, 1987).

546. Rutstein, D.D., Berenberg, W., Chalmers, T.C., et al., "Measuring the Quality of Medical Care: A Clinical Method," *New England Journal of Medicine* 294(11):582-588, 1976.

547. Safran, C., Center for Clinical Computing, Beth Israel Hospital, Boston, MA, presentation at the American Public Health Association Meeting, New Orleans, LA, Oct. 18, 1987.

548. Saks, M.J., "In Search of the 'Lawsuit Crisis'," *Law, Medicine and Health Care* 14(2):77-82, 1987.

549. Samet, J., Hunt, W.C., Key, C., et al., "Choice of Cancer Therapy Varies With Age of Patient," *Journal of the American Medical Association* 255:3385, 1986.

550. Sanazaro, P.J., "Measurement of Physicians' Performance Using Existing Techniques," *Western Journal of Medicine* 133(7):81-88, 1980.

551. Sanazaro, P.J., "Predictive Validity and Measurement of Clinical Performance," *Evaluating the Skills of Medical Specialists* (Chicago, IL: American Board of Medical Specialties, 1983).

552. Sanazaro, P.J., and Worth, R.M., "Measuring Clinical Performance of Individual Internists in Office and Clinical Practice," *Medical Care* 23(9):1097-1114, 1985.

553. Saxe, L., *The Efficacy and Cost Effectiveness of Psychotherapy* (Health Technology Case Study #18), prepared for the Office of Technology Assessment, U.S. Congress, OTA-BP-H-9(18) (Washington, DC: U.S. Government Printing Office, October 1980).

554. Saxe, L., and Fine, M., *Social Experiments* (Beverly Hills, CA: Sage Publishing Co., 1981).

555. Schimmel, E.M., "The Hazards of Hospitalization," *Annals of Internal Medicine* 60:100, 1964.

556. Schoen, C., Research Director, Service Employees International Union, Washington, DC, personal communication, February 1988.

557. Schroeder, S.A., and Donaldson, M.S., "The Feasibility of an Outcome Approach to Quality Assurance—A Report From One HMO," *Medical Care* 14(1):49-56, 1976.

558. Schulman, B.A., "Active Patient Orientation and Outcomes in Hypertensive Treatment," *Medical Care* 17(3):267-280, 1979.

559. Schyve, P.M., "Standards Development Model," prepared for the Joint Commission on

Accreditation of Healthcare Organizations, Chicago, IL, 1987.

560. Schyve, P.M., Director, Department of Standards, Joint Commission on Accreditation of Healthcare Organizations, Chicago, IL, personal communication, Feb. 19, 1987.

561. Scott, W.R., and Flood, A.B., "Costs and Quality of Hospital Care: A Review of the Literature," *Hospital Privileges and Specialty Medicine* (Chicago, IL: American Board of Medical Specialties, 1986).

562. Scott, W.R., Flood, A.B., and Ewy, W., "Organizational Determinants of Services, Quality and the Cost of Care in Hospitals," *Hospital Structure and Performance*, A.B. Flood and W.R. Scott (eds.) (Baltimore, MD: Johns Hopkins University Press, 1987).

563. Sehnert, K.W., and Eisenberg, H., *How To Be Your Own Doctor (Sometimes)* (New York, NY: Grossett & Dunlap, 1975).

564. Shackford, S., Hollingsworth-Fridlund, P., McArdle, M., et al., "Assuring Quality in a Trauma System—The Medical Audit Committee: Composition, Cost, and Results," *Journal of Trauma* 27(8):866-875, 1987.

565. Shank, R., and Abelson, R., *Scripts, Plans, Goals, and Understanding* (Hillsdale, NJ: Lawrence Erlbaum Association, 1977).

566. Shapiro, M.F., Ware, J.E., Jr., and Sherbourne, C.D., "Effects of Cost Sharing on Seeking Care for Serious and Minor Symptoms: Results of a Randomized Controlled Trial," *Annals of Internal Medicine* 104:246-251, 1986.

567. Sheps, S., and Robertson, A., "Evaluation of Primary Care in a Community Clinic by Means of Explicit Process Criteria," *Canadian Medical Association Journal* 131:881-886, 1984.

568. Shimberg, B., "The Relationship Among Accreditation, Certification and Licensure," *Federation Bulletin* 71(4):99-116, 1984.

569. Shorr, G.I., and Nutting, P.A., "A Population-Based Assessment of the Continuity of Ambulatory Care," *Medical Care* 15(6):455-464, 1977.

570. Shortell, S.M., Becker, S.W., and Neuhauser, D., "The Effects of Management Practices on Hospital Efficiency and Quality of Care," *Organizational Research in Hospitals*, S.M. Shortell and M. Brown (eds.) (Chicago, IL: Blue Cross Association, 1976).

571. Shortell, S.M., and LoGerfo, J.P., "Hospital Medical Staff Organization and Quality of Care: Results for Myocardial Infarction and Appendectomy," *Medical Care* 19(10):1041-1053, 1981.

572. Shortell, S.M., Longo, D.R., Hughes, E.F.X., et al., "Assessing the Reliability of Selected Joint Commission Standards," prepared for the Joint Commission on the Accreditation of Hospitals, Chicago, IL, September 1986.

573. Showstack, J.A, Rosenfeld, K.E., Garnick, D.W., et al., "Association of Volume With Outcome of Coronary Artery Bypass Graft Surgery," *Journal of the American Medical Association* 257(6):785-789, 1987.

574. Showstack, J.A, Rosenfeld, K.E., Garnick, D.W., et al., Institute for Health Policy Studies, University of California, unpublished data, San Francisco, CA, 1987.

575. Shugan, S.M., "Price-Quality Relationships," *Advances in Consumer Research, Vol. 11*, T. Kinnear (ed.) (Ann Arbor, MI: Association for Consumer Research, 1984).

576. Sibley, J.C., Spitzer, W.O., Rudnick, A.V., et al., "Quality-of-Care Appraisal in Primary Care: A Quantitative Method," *Annals of Internal Medicine* 83:46-52, 1974.

577. Simon, H.A., "How Big Is a Chunk," *Science* 183(4124):482-488, 1974.

578. Simpson, J.B., "PRO Release of Outcome Data: Legal Issues," Western Consortium for the Health Professions, Inc., San Francisco, CA, 1987.

579. Simpson, J.B., "Release of Physician-Specific Quality of Care Information: Legal Issues," contractor document prepared for the Office of Technology Assessment, U.S. Congress, Washington, DC, January 1988.

580. Sinks, L.F., Chief, Cancer Centers Branch, Division of Cancer Prevention and Control, National Cancer Institute, National Institutes of Health, Public Health Service, U.S. Department of Health and Human Services, Bethesda, MD, personal communication, January 1988.

581. Slater, C., School of Public Health, University of Texas Health Science Center, Houston, TX, personal communication, August 1987.

582. Sloan, F.A., Perrin, J.M., and Valvona, F., "In-Hospital Mortality of Surgical Patients: Is There an Empiric Basis for Standard Setting?" *Surgery* 99(4):446-453, 1986.

583. Smith, C.K., Polis, E., and Hadac, R.R., "Characteristics of the Initial Medical Interview Associated With Patient Satisfaction and Understanding," *Journal of Family Practice* 12:283-288, 1981.

584. Smith, M.L., Glass, G.V., and Miller, T.I., *The Benefits of Psychotherapy* (Baltimore, MD: Johns Hopkins University Press, 1980).

585. Souelem, O., "Mental Patients' Attitudes Toward Mental Hospitals," *Journal of Clinical Psychology* 7:329-337, 1955.

586. Sox, H.C., Jr., Margulies, I., and Sox, C.H., "Psychologically Mediated Effects of Diagnostic Tests," *Annals of Internal Medicine* 95:680-685, 1981.

587. Spengler, R.F., and Greenough, W.B., "Hospital Costs and Mortality Attributed to Nosocomial Bacteremias," *Journal of the American Medical Association* 240(22):2455-2458, 1978.

588. Stanford Center for Health Care Research, *Study of Institutional Differences in Postoperative Mortality: A Report to the National Academy of Sciences, National Research Council*, NTIS Pub. No. PB 250 940 (Springfield, VA: National Technical Information Service, 1974).

589. Stanford Center for Health Care Research, "Comparison of Hospitals With Regard to Outcomes of Surgery," *Health Services Research* 11(2):112-127, 1976.

590. Stanford Center for Health Care Research, *Impact of Hospital Characteristics on Surgical Outcomes and Length of Stay*, NTIS Pub. No. PB81-233165 (Springfield, VA: National Technical Information Service, 1978).

591. State of New York, Department of Health, "Incident Reporting for Hospitals," Health Facilities Series H-61, Albany, NY, Sept. 30, 1985.

592. State of New York, Department of Health, "Incident Reporting for Hospitals," Health Facilities Series H-85, Albany, NY, Dec. 29, 1986.

593. State of New York, Department of Health, "Hospital Incident Reporting Program, 1986 Annual Report," Albany, NY, Mar. 17, 1987.

594. State of New York, Department of Health, *Investigation of Quality of Care Problems Associated With In-Hospital Mortality in New York State: An Identification of Critical Case Characteristics for Targeting Purposes* (Albany, NY: December 1987).

595. Steel, K. Gertman, P.M., Crescenzi, C., et al., "Iatrogenic Illness on a General Medical Service at a University Hospital," *New England Journal of Medicine* 304(11):638-642, 1981.

596. Stein, R.E.K., Gortmaker, S.L., Perrin, E.C., et al., "Severity of Illness: Concepts and Measurements," *Lancet* 2(8574):1506-1509, 1987.

597. Steinbrook, R., "Hospital Death Rates: A Wide Variance," *Los Angeles Times*, p. 1, June 15, 1987.

598. Steinbrook, R., "Care for Newborns Varies, Studies of Hospitals Show," *Los Angeles Times*, p.1, Nov. 9, 1987.

599. Stewart, D.W., Hickson, G.B., Ratneshwar, S., et al., "Information Search and Decision Strategies Among Health-Care Consumers," *Advances in* Consumer Research 12:252-257, 1985.

600. Stewart, M.A., "What Is a Successful Doctor-Patient Interview? A Study of Interactions," *Social Science and Medicine* 19:167-175, 1984.

601. Stewart, M.A., McWhinney, I.R., and Buck, C.W., "The Doctor/Patient Relationship and its Effect Upon Outcome," *Journal of the Royal College of General Practitioners* 29(199):77-81, 1979.

602. Stiles, W.B., Putnam, S.M., James, S.A., et al., "Dimensions of Patient and Physician Roles in Medical Screening Interviews," *Social Science and Medicine* 13A:335-341, 1979.

603. Stratmann, W.C., "A Study of Consumer Attitudes About Health Care: The Delivery of Ambulatory Services," *Medical Care* 13(7):537-548, 1975.

604. Strauss, M.J., Conrad, D., LoGerfo, J.P., et al., "Cost and Outcome of Care for Patients With Chronic Obstructive Lung Disease," *Medical Care* 24(10):915-924, 1986.

605. Sullivan, R., "Cuomo Plan for Physician Reviews Gains Support," *The New York Times*, p. 29, June 28, 1987.

606. Sulvetta, M.B., and Swartz, K, "The Uninsured and Uncompensated Care: A Chartbook," prepared for the National Health Policy Forum, George Washington University, Washington, DC, June 1986.

607. Summer, S.J., "Maryland's Experiment With Quality Measures," *Business and Health* 5(1):14-16, 1987.

608. Sunshine, L., and Wright, J.W., *The Best Hospitals in America* (New York, NY: H. Holt & Co., 1987).

609. Swartz, T.A., and Stephens, N, "Information Search for Services: The Maturity Segment," *Advances in Consumer Research, Vol. 11*, T.C. Kinnear (ed.) (Provo, UT: Association for Consumer Research, 1984).

610. SysteMetrics, "Report on Available State-Specific Databases," contractor document prepared for the Office of Technology Assessment, U.S. Congress, Washington, DC, September 1987.

611. Tancredi, L.R., "Compensating for Medical Injuries: Is There an Effective Alternative to the Tort System of Medical Malpractice," *New York State Journal of Medicine* 86(7):370-376, 1986.

612. Templeton, B., Vivian, E.F., Bunce, J.V., et al., "Use of a Medical Record Audit in Assessing Pediatric Resident Performance: A Preliminary

Report," *Evaluating the Skills of Medical Specialists* (Chicago, IL: American Board of Medical Specialties, 1983).

613. Terry, M.K., Director, Division of Review Programs, U.S. Department of Health and Human Services, Bethesda, MD, personal communication, June 10, 1987.

614. Thomas, J.W., Ashcraft, M.L.F., and Zimmerman, J., *An Evaluation of Alternative Severity of Illness Measures for Use by University Hospitals* (Ann Arbor, MI: University of Michigan Department of Health Services Management and Policy, Dec. 29, 1986).

615. Thompson, H.C., and Osborne, C.E., "Office Records in the Evaluation of Quality of Care," *Medical Care* 14:294-314, 1976.

616. Thompson, R.L., "Surveillance and Reporting of Nosocomial Infections," *Prevention and Control of Nosocomial Infections*, R.P. Wenzel (ed.) (Baltimore, MD: Williams & Wilkins, 1987).

617. Thorelli, H.B., and Engledow, J.L., "Information Seekers and Information Systems: A Policy Perspective," *Journal of Marketing* 44(2):9-24, 1980.

618. Thorndike, R.L., *Applied Psychometrics* (Boston, MA: Houghton Mifflin, 1982).

619. Tversky, A., and Kahneman, D., "Choices, Values and Frames," *American Psychologist* 39(4):341-350, 1984.

620. U.S. Congress, General Accounting Office, *Medicare Past Overuse of Intensive Care Services Inflates Hospital Payments* (Washington, DC: U.S. Government Printing Office, March 1986).

621. U.S. Congress, General Accounting Office, *States Assume Leadership Role in Providing Emergency Medical Services*, HRD-86-132 (Washington, DC: September 1986).

622. U.S. Congress, General Accounting Office, *Medicare: Preliminary Strategies for Assessing Quality of Care*, GAO/PEMD-87-15BR (Washington, DC: U.S. Government Printing Office, 1987).

623. U.S. Congress, General Accounting Office, *Medical Malpractice: Characteristics of Claims Closed in 1984*, GAO/HRD-87-55 (Washington, DC: U.S. Government Printing Office, Apr. 22, 1987).

624. U.S. Congress, General Accounting Office, *VA's Patient Injury Control Program Not Effective*, GAO/HRD-87-49 (Washington, DC: U.S. Government Printing Office, May 1987).

625. U.S. Congress, General Accounting Office, *Medical Malpractice: A Framework for Action*, GAO/HRD-87-73 (Washington, DC: U.S. Government Printing Office, May 1987).

626. U.S. Congress, General Accounting Office, "Medicare: Improved Patient Outcome Analyses Could Enhance Quality Assessment," draft report, Washington, DC, forthcoming.

627. U.S. Congress, Library of Congress, Congressional Research Service, *The Peer Review Organization (PRO) Program*, report to the Senate Committee on Finance, U.S. Congress, 87-258 EPW (Washington, DC: U.S. Government Printing Office, 1987).

628. U.S. Congress, Office of Technology Assessment, *Assessing the Efficacy and Safety of Medical Technologies*, OTA-H-75 (Washington, DC: U.S. Government Printing Office, September 1978).

629. U.S. Congress, Office of Technology Assessment, *Diagnosis-Related Groups (DRGs) and the Medicare Program: Implications for Medical Technology*, OTA-TM-H-17 (Washington, DC: U.S. Government Printing Office, July 1983).

630. U.S. Congress, Office of Technology Assessment, *Medicare's Prospective Payment System: Strategies for Evaluating Cost, Quality, and Medical Technology*, OTA-H-262 (Washington, DC: U.S. Government Printing Office, October 1985).

631. U.S. Congress, Office of Technology Assessment, *Payment for Physician Services: Strategies for Medicare*, OTA-H-294 (Washington, DC: U.S. Government Printing Office, February 1986).

632. U.S. Congress, Office of Technology Assessment, *Life-Sustaining Technologies and the Elderly*, OTA-BA-306 (Washington, DC: U.S. Government Printing Office, July 1987).

633. U.S. Department of Agriculture and U.S. Department of Health and Human Services, "Design and Evaluation of Nutrition Label Formats: Information Kit," Washington, DC, Dec. 2, 1982.

634. U.S. Department of Commerce, Bureau of the Census, "Detailed Population Characteristics, United States Summary Section A," Washington, DC, March 1984.

635. U.S. Department of Commerce, Bureau of the Census, "Current Population Survey," unpublished data, Washington, DC, March 1986.

636. U.S. Department of Health and Human Services, "Report of the Task Force on Medical Liability and Malpractice," Washington, DC, August 1987.

637. U.S. Department of Health and Human Services, Health Care Financing Administration, *SuperPRO Request for Proposal* (Baltimore, MD: 1985).

638. U.S. Department of Health and Human Services, Health Care Financing Administration, "Peer Review Organization Manual," HCFA Pub. No. 19, Transmittal No. 3, Baltimore, MD, August 1985.

639. U.S. Department of Health and Human Services, Health Care Financing Administration, *Medicare Annual Report Fiscal Year 1983* (Washington, DC: U.S. Government Printing Office, 1986).

640. U.S. Department of Health and Human Services, Health Care Financing Administration, "Medicare Hospital Mortality Information 1984," Washington, DC, Mar. 10, 1986.

641. U.S. Department of Health and Human Services, Health Care Financing Administration, *Medicare/Medicaid Hospital Survey Report*, Form HCFA-1537 (Baltimore, MD: April 1986).

642. U.S. Department of Health and Human Services, Health Care Financing Administration, "Appendix A: Interpretive Guidelines and Survey Procedures—Hospitals," *State Operations Manual, Provider Certification*, HCFA Pub. 7 (Baltimore, MD: September 1986).

643. U.S. Department of Health and Human Services, Health Care Financing Administration, "Medicare Program: Selective Performance Information on Hospitals Providing Care to Medicare Beneficiaries," *Federal Register* 32(158):30741-30745, 1987.

644. U.S. Department of Health and Human Services, Health Care Financing Administration, *Scope of Work for Contract To Perform HMO/CMP Review* (Baltimore, MD: 1987).

645. U.S. Department of Health and Human Services, Health Care Financing Administration, "Peer Review Organization Manual," HCFA Pub. No. 19, Transmittal No. 15, Baltimore, MD, May 1987.

646. U.S. Department of Health and Human Services, Health Care Financing Administration, "National Surveyor Characteristics Training Questionnaire Report: Highest Educational Attainment," Baltimore, MD, June 16, 1987.

647. U.S. Department of Health and Human Services, Health Care Financing Administration, *Medicare Hospital Mortality Information, 1986* (Washington, DC: U.S. Government Printing Office, Dec. 17, 1987).

648. U.S. Department of Health and Human Services, Health Care Financing Administration, HHS News, press release and untitled questions and answers accompanying *Medicare Hospital Mortality Information 1986*, Washington, DC, Dec. 17, 1987.

649. U.S. Department of Health and Human Services, Health Care Financing Administration, "Individual Hospital Facilities Profiles," Baltimore, MD, 1988.

650. U.S. Department of Health and Human Services, Health Care Financing Administration, "Quarterly Comparison of State, Regional, and National Deficiency Patterns in Prefix Number Sequence," Baltimore, MD, Jan. 2, 1988.

651. U.S. Department of Health and Human Services, Health Care Financing Administration, Health Standards and Quality Bureau, *1986-1988 PRO Scope of Work* (Baltimore, MD: Nov. 4, 1985).

652. U.S. Department of Health and Human Services, Health Care Financing Administration, Health Standards and Quality Bureau, *1988-1990 PRO Scope of Work* (Baltimore, MD: Apr. 1, 1988).

653. U.S. Department of Health and Human Services, Health Care Financing Administration, Health Standards and Quality Bureau, Office of Medical Review, "Efficiency and Effectiveness of the Generic Quality Screens," Baltimore, MD, May 1, 1987.

654. U.S. Department of Health and Human Services, Health Care Financing Administration, Health Standards and Quality Bureau, Office of Medical Review, "Comparison of PRO Versus SuperPRO Change Rates," Baltimore, MD, Oct. 30, 1987.

655. U.S. Department of Health and Human Services, Health Care Financing Administration, Health Standards and Quality Bureau, Office of Medical Review, "HCFA's Generic Quality Screens," Baltimore, MD, undated.

656. U.S. Department of Health and Human Services, Health Care Financing Administration, Health Standards and Quality Bureau, Office of Medical Review, Division of Program Operations, "Corrective Action Plan Status Report," Baltimore, MD, November 1987.

657. U.S. Department of Health and Human Services, Health Information Policy Council, "1984 Revision of the Uniform Hospital Discharge Data Set," *Coding Clinic* 2(4):3-8, 1985.

658. U.S. Department of Health and Human Services, Office of the Inspector General, Office of Analysis and Inspections, "Medical Licensure and Discipline: An Overview," Control No. P-01-86-0064, Washington, DC, June 1986.

659. U.S. Department of Health and Human Services, Office of the Inspector General, Office of Analysis and Inspections, *Peer Review Organization (PRO) Program Performance* (Washington, DC: April 1987).

660. U.S. Department of Health and Human Services, Office of the Inspector General, Office of Analysis and Inspections, *National DRG Validation Study: Special Report on Coding Accuracy* (Washington, DC: February 1988).

660a. U.S. Department of Health and Human Services, Office of the Inspector General, Office of Analysis and Inspections, *National DRG Validation Study: Special Report on Premature Discharges* (Washington, DC: February 1988).

661. U.S. Department of Health and Human Services, Office of the Inspector General, Office of Analysis and Inspections, "The Utilization and Quality Control Peer Review Organization (PRO) Program," draft report, Control No. OAI-01-88-00570, Washington, DC, February 1988.

662. U.S. Department of Health and Human Services, Public Health Service, *Promoting Health/Preventing Disease: Objectives for the Nation* (Washington, DC: 1980).

663. U.S. Department of Health and Human Services, Public Health Service, National Center for Health Statistics, "Annual Summary of Births, Deaths, Marriages, and Divorces," *Monthly Vital Statistics Report* 34:13 (DHHS Pub. No. [PHS] 86-1120) (Hyattsville, MD: 1986).

664. U.S. Department of Health and Human Services, Public Health Service, National Center for Health Statistics, "1985 Summary: National Hospital Discharge Survey," *NCHS Advance Data*, No. 127, Hyattsville, MD, Sept. 25, 1986.

665. U.S. Department of Health and Human Services, Public Health Service, National Center for Health Statistics, unpublished data from the National Health Interview Survey, Hyattsville, MD, Nov. 7, 1986.

666. U.S. Department of Health and Human Services, Public Health Service, National Center for Health Statistics, unpublished data from the National Ambulatory Medical Care Survey, Hyattsville, MD, Nov. 17, 1986 and Jan. 16, 1987.

667. U.S. Department of Health and Human Services, Public Health Service, National Center for Health Statistics, *Utilization of Short-Stay Hospitals, United States, 1985, Annual Summary*, DHHS Pub. No. [PHS] 87-1752, Series 13, No. 91 (Hyattsville, MD: 1987).

668. U.S. Department of Health and Human Services, Public Health Service, National Institutes of Health, National Cancer Institute, "Guidelines for Recognition of a Cancer Center as Comprehensive," Bethesda, MD, June 14, 1985.

669. U.S. Executive Office of the President, Office of Management and Budget, *Budget of the United States Government* (Washington, DC: U.S. Government Printing Office, 1988).

670. U.S. Government Printing Office, Superintendent of Documents, unpublished data, Washington, DC, May 1988.

671. VanDoren, D.C., and Smith, L.W., "Physician Marketing in the Restructured Medical Services Field," *Journal of Health Care Marketing* 7(3):7-14, 1987.

672. Van Veen Daigle, A., "The AMA Physician Masterfile: A Credentialing Resource," *Resident & Staff Physician* 32(5):183-187, 1986.

673. Van Veen Daigle, A., Director, Data Planning and Evaluation, American Medical Association, Chicago, IL, personal communication, Feb. 19, 1987.

674. *Varandani v. Bowen*, 824 F.2d 307 (4th Cir. 1987).

675. Venezian, E.C. (Rutgers University), Nye, B.F. (the MAC Group), and Hofflander, A.E. (University of California at Los Angeles), "The Distribution of Claims for Professional Malpractice—Some Statistical and Public Policy Aspects," 1986.

676. Veterans Administration Cooperative Study Group on Antihypertensive Agents, "Effects of Treatment on Morbidity in Hypertension, Part I: Results in Patients With Diastolic Blood Pressure Averaging 115 Through 129 mmHg," *Journal of the American Medical Association* 202:1028-1034, 1967.

677. Veterans Administration Cooperative Study Group on Antihypertensive Agents, "Effects of Treatment on Morbidity in Hypertension, Part II: Results in Patients With Diastolic Blood Pressure Averaging 90 to 114 mmHg," *Journal of the American Medical Association* 213:1143-1152, 1970.

678. Vickory, M.D., and Fries, J.F., *Take Care of Yourself: A Consumer's Guide to Medical Care* (Reading, MA: Addison-Wesley, 1976).

679. Wales, J., Kane, R., Robbins, S., et al., "UCLA Hospice Evaluation Study: Methodology and Instrumentation," *Medical Care* 21:734-744, 1983.

680. Walker, A.H., and Restuccia, J.D., "Obtaining Information on Patient Satisfaction With Hos-

pital Care: Mail vs. Telephone," *Health Services Research* 19:291-306, 1984.

681. Ward, T.F., "Credentialling and Privilege Reviews Need Objective Criteria," *Medical Staff Affairs* 2:34-37, 1987.

682. Ware, J.E., Jr., "Effects of Acquiescent Response Set on Patient Satisfaction Ratings," *Medical Care* 16(4):327-336, 1978.

683. Ware, J.E., Jr., Senior Research Psychologist, the RAND Corp., unpublished data, Santa Monica, CA, March 1988.

684. Ware, J.E., Jr., Curbow, B., Davies, A.R., et al., *Medicaid Satisfaction Surveys, 1977-1980: A Report of the Prepaid Health Research, Evaluation, and Demonstration Project* (Sacramento, CA: California State Department of Health Services, 1981).

685. Ware, J.E., Jr., and Davies, A.R., "Behavioral Consequences of Consumer Dissatisfaction With Medical Care," *Evaluation and Program Planning* 6:291-297, 1983.

686. Ware, J.E., Jr., Davies, A.R., and Rubin, H.R., "The Suitability of Consumers' Assessments of Physician and Hospital Performance as Indicators of the Quality of Care," contractor document prepared for the Office of Technology Assessment, U.S. Congress, Washington, DC, 1987.

687. Ware, J.E., Jr., Davies-Avery, A., and Stewart, A.L., "The Measurement and Meaning of Patient Satisfaction," *Health & Medical Care Services Review* 1(1):3-15, 1978.

688. Ware, J.E., Jr., and Hays, R., "Methods for Measuring Patient Satisfaction With Specific Medical Encounters," *Medical Care* 26(4):393-402, 1988.

689. Ware, J.E., Jr., Kane, R.L., Davies, A.R., et al., "The Patient's Role in Assessing the Quality of the Medical Care Process," in progress.

690. Ware, J.E., Jr., and Snyder, M.K., "Dimensions of Patient Attitudes Regarding Doctors and Medical Care Services," *Medical Care* 13:669-682, 1975.

691. Ware, J.E., Jr., Snyder, M.K., and Wright, W.R., *Development and Validation of Scales To Measure Patient Satisfaction With Health Care Services, Final Report, Vol. I, Part B: Results Regarding Scales Constructed From the Patient Satisfaction Questionnaire and Measures of Other Health Care Perceptions*, NTIS Pub. No. PB 288-330 (Springfield, VA: National Technical Information Service, 1976).

692. Ware, J.E., Jr., Wright, W.R., Snyder, M.K., et al., "Consumer Perceptions of Health Care Services: Implications for Academic Medicine," *Journal of Medical Education* 50:839-848, 1975.

693. Washington Consumers' Checkbook, "Hospitals," *Washington Consumers' Checkbook* 6(3):13-27, 1987.

694. Watry, A., Executive Director, Georgia Composite State Board of Medical Examiners, Atlanta, GA, personal communication, Oct. 1, 1987.

695. Weinberger, M., Greene, J.Y., and Mamlin, J.J., "The Impact of Clinical Encounter Events on Patient and Physician Satisfaction," *Social Science and Medicine* 15E:239-244, 1981.

696. Wennberg, J.E., "The Paradox of Appropriate Care," *Journal of the American Medical Association* 258(18):2568-2569, 1987.

697. Wennberg, J.E., Roos, N., Sola, L., et al., "Use of Claims Data Systems To Evaluate Health Care Outcomes: Mortality and Reoperation Following Prostatectomy," *Journal of the American Medical Association* 257(7):933-936, 1987.

698. Wenzel, R.P., "Nosocomial Infections, Diagnosis-Related Groups, and Study on the Efficacy of Nosocomial Infection Control/Economic Implications for Hospitals Under the Prospective Payment System," *American Journal of Medicine* 78(Suppl. 6B):3-7, 1985.

699. Wenzel, R.P., Osterman, C.A., Donowitz, L.G., et al., "Identification of Procedure-Related Nosocomial Infections in High-Risk Patients," *Review of Infectious Diseases* 3(4):701-707, 1981.

700. West, J.G., Cales, R.H., Gazzaniga, A.G. (eds.), *Trauma Care Systems: Clinical, Financial, and Political Considerations* (New York, NY: Praeger, 1983).

700a. Wetle, T., "Age as a Risk Factor for Inadequate Treatment," *Journal of the American Medical Association* 258(4):516, 1987.

701. Whiting-O'Keefe, Q.E., Henke, C., and Simborg, D.W., "Choosing the Correct Unit of Analysis in Medical Care Experiments," *Medical Care* 22(12):1101-1114, 1984.

702. Williams, R.L., "Measuring the Effectiveness of Perinatal Medical Care," *Medical Care* 17(2):95-109, 1979.

703. Williamson, J.W., "Validation By Performance Measures," *Conference on Extending the Validity of Certification, 1976* (Chicago, IL: American Board of Medical Specialties, 1976).

704. Williamson, J.W., *Assessing and Improving Health Care Outcomes: The Health Accounting Approach to Quality Assurance* (Cambridge, MA: Ballinger Publishing Co., 1978).

705. Williamson, J.W., "Improving Content Validity of Certification Procedures by Defining Competence in Specialty Practice: Directions, Resources, and Getting Started," *Evaluating the Skills of Medical Specialists* (Chicago, IL: American Board of Medical Specialties, 1983).

706. Wilson, M., Health Manpower Education Specialist, Division of Medicine, Health Resources and Services Administration, Rockville, MD, personal communication, Sept. 22, 1987.

707. Wimsatt, W.D., "Robustness, Reliability and Overdetermination," *Scientific Inquiry and the Social Sciences*, M.B. Brewer and B.E. Collins (eds.) (San Francisco, CA: Jossey-Bass, 1981).

708. Winickoff, R.N., Wilner, S., Neisuler, R., et al., "Limitations of Provider Interventions in Hypertension Quality Assurance," *American* Journal of Public Health 75(1):43-46, 1985.

709. Wiseman, M., Director, Practice Management Department, American College of Emergency Physicians, Dallas, TX, personal communication, Mar. 15, 1988.

710. Wolf, F.M., *Meta-Analysis: Quantitative Methods for Research Synthesis* (Beverly Hills, CA: Sage Publishing Co., 1986).

711. Wolf, R.N., "Validity—Basic Ideas and Their Use," *Conference on Extending the Validity of Certification, 1976* (Chicago, IL: American Board of Medical Specialties, 1976).

712. Wolfe, S.M., Director, Public Citizen Health Research Group, testimony presented before the New York City Council, New York, NY, Jan. 21, 1986.

713. Wolfe, S.M., Director, Public Citizen Health Research Group, testimony at hearing on Medical Peer Review Organization Program before the Subcommittee on Health and the Environment, Committee on Energy and Commerce, House of Representatives, U.S. Congress, Oct. 26, 1987.

714. Wolinsky, F.D., and Steiber, S.A., "Salient Issues in Choosing a New Doctor," *Social Science and Medicine* 16(7):759-767, 1982.

715. Wood, D., Division of Information Analysis, Bureau of Data Management and Strategy, Health Care Financing Administration, U.S. Department of Health and Human Services, Baltimore, MD, personal communication, Jan. 19, 1988.

716. Woodward, C.A., McConvey, G.A., Neufeld, V., et al., "Measurement of Physician Performance by Standardized Patients," *Medical Care* 23(8):1019-1027, 1985.

717. Working Group on Hypertension in the Elderly, "Statement on Hypertension in the Elderly," *Journal of the American Medical Association* 256(1):70-74, 1986.

718. World Health Organization, Constitution, 1948, as cited in J.S. Mausner and A.K. Bahn, *Epidemiology: An Introductory Text* (Philadelphia, PA: W.B. Saunders Co., 1974).

719. Wotruba, T.R., Haas, R.W., and Oulhen, H., "Marketing Factors Affecting Physician Choice as Related to Consumers' Extent of Use and Predisposition Toward Use of Physician Services," *Journal of Health Care Marketing* 5(4):7-17, 1985.

720. Yessian, M.R., Ph.D, Regional Inspector General, Office of Analysis and Inspections, Office of the Inspector General, Department of Health and Human Services, Boston, MA, personal communication, July 13, 1987.

721. York, M., and Moore, T.J., "Bypass Surgery: Is it Regulated Enough," *Lexington Herald-Leader*, p. A1, Sept. 22, 1986.

722. Zyzanski, S.J., Hulka, B.S., and Cassel, J.C., "Scale for the Measurement of 'Satisfaction' With Medical Care," *Medical Care* 12:611-620, 1974.

Superintendent of Documents Publication Order Form

Order Processing Code
***6418**

Charge your order.
It's easy! MasterCard VISA CHOICE

☐ **YES,** please send me the following indicated publications:

The Quality of Medical Care: Information for Consumers.
GPO stock number 052-003-01114-0; price $14.00.

1. The total cost of my order is $_______ (International customers please add an additional 25%.) All prices include regular domestic postage and handling and are good through 11/88. After this date, please call Order and Information Desk at 202-783-3238 to verify prices.

Please Type or Print

2. ____________________________________
(Company or personal name)

____
(Additional address/attention line)

____
(Street address)

____
(City, State, ZIP Code)

(_______) ____________________________
(Daytime phone including area code)

3. Please choose method of payment:

☐ Check payable to the Superintendent of Documents
☐ GPO Deposit Account ☐☐☐☐☐☐☐–☐
☐ VISA, CHOICE or MasterCard Account
☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐

____
(Credit card expiration date)

Thank you for your order!

____
(Signature) 6/88

4. Mail To: Superintendent of Documents, Government Printing Office, Washington, D.C. 20402-9325

Office of Technology Assessment

The Office of Technology Assessment (OTA) was created in 1972 as an analytical arm of Congress. OTA's basic function is to help legislative policymakers anticipate and plan for the consequences of technological changes and to examine the many ways, expected and unexpected, in which technology affects people's lives. The assessment of technology calls for exploration of the physical, biological, economic, social, and political impacts that can result from applications of scientific knowledge. OTA provides Congress with independent and timely information about the potential effects—both beneficial and harmful—of technological applications.

Requests for studies are made by chairmen of standing committees of the House of Representatives or Senate; by the Technology Assessment Board, the governing body of OTA; or by the Director of OTA in consultation with the Board.

The Technology Assessment Board is composed of six members of the House, six members of the Senate, and the OTA Director, who is a nonvoting member.

OTA has studies under way in nine program areas: energy and materials; industry, technology, and employment; international security and commerce; biological applications; food and renewable resources; health; communication and information technologies; oceans and environment; and science, education, and transportation.

OTA-H-386

JUNE 1988